TIGHT BINDING BOOK

PHYSICS

A Supplementary Textbook
for the use of Scholarship Candidates
and University Students

BY

G. T. P. TARRANT
M. A., Ph. D..

VOLUME ONE

MECHANICS
PROPERTIES OF MATTER
LIGHT & HEAT

PREFACE

THIS book is designed for students who have already reached the Higher Certificate standard in physics and are proceeding to university scholarships or degrees. It is to be regarded rather as a supplement to existing textbooks than as an addition to their number, and takes its origin from my teaching experience at Cambridge and Manchester. Any points that students found difficult, or of which a clearer treatment than usual could be developed, were noted down and repeatedly considered with successive classes. Certain other topics that seemed to present features of special interest or importance were also recorded, and from the data thus accumulated the book gradually took shape.

Although the several sections of the book are not necessarily interconnected, each section is continuous within itself and includes all that a scholarship or degree candidate is likely to want—apart, of course, from such readily assimilable book-work as he may easily obtain from his ordinary textbook. The whole of the matter has been carefully sifted in order that the essentials of each thesis shall not be obscured by un-nec ssary accretions likely to cause confusion; and it may not be amiss to remark that the book represents such treatment as I have found in practice to be of both immediate and lasting help to the student. In a further volume I hope to deal on similar lines with electricity, magnetism, and modern physics.

I am grateful to the following firms for permission to reproduce diagrams and tables: Messrs. W. Edwards & Co., London, for Figure 27; the Cambridge University Press for the three diagrams of Figure 31, from *Experimental Elasticity* by G. F. C. Searle; the Cambridge University Press and the International Commission on Illumination for

vi

tables which appear in the chapters on Colour; and Messrs. Longmans, Green & Co. for Figure 142, from *Treatise on Light* by R. A. Houstoun.

I feel very greatly indebted to all my students, from whom I have learnt much, and, in particular, to the scholarship candidates at Manchester who have read—and criticized—the manuscript.

My thanks are also due to Dr. N. Feather for his help in connection with certain difficult points, to Dr. H. Lowry for reading the manuscript of the chapter on Colour, and to my wife for much help, advice, and forbearance.

G. T. P. T.

CONTENTS

CHAPTER IV

PRISMS, LENSES, AND TELESCOPIC SYSTEMS

CONTENTS

CHAPTER V

THE NATURE AND POLARIZATION OF LIGHT

CHAPTER VI

PHOTOMETRY

CHAPTER VII
COLOUR

CHAPTER VIII
INTERFERENCE

CHAPTER IX
DIFFRACTION

CHAPTER X

MICHELSON'S METHOD FOR THE DETERMINATION OF THE VELOCITY OF LIGHT . . . . .

CHAPTER XI

THE GAS SCALE AND THERMODYNAMICS

CONTENTS

CHAPTER XII

LIQUEFACTION AND LOW TEMPERATURES

CHAPTER I

MECHANICS

§ 1. Straight Line and Angular Mechanics. § 2. Simple Harmonic Motion. § 3. Centrifugal Force. § 4. The Wave Equation. § 5. Stationary Waves. Mathematical Treatment. § 6. Stationary Waves. Physical Treatment. § 7. Errors. § 8. The Combination of Errors. § 9. Independent Measurements.

§ 1. Straight Line and Angular Mechanics

A knowledge of the methods for proving basic formulae is probably of less importance in the study of mechanics than is the development of the technique of applying these formulae to new problems. In consequence it often happens that students concentrate entirely on the problems and largely forget the methods by which the fundamental formulae were obtained, so that it seems valuable to summarize these proofs.

The quantities and units which are involved in the study of mechanics are all derived ultimately from three fundamentals, viz., *mass*, *length*, and *time*. These are taken as fundamental because they are the easiest to measure and to reproduce, and have internationally agreed values.

Velocity is the name given to the ratio $\dfrac{\text{distance}}{\text{time}}$; hence, in the limit,

$$v = \frac{ds}{dt}.$$

Acceleration is the name given to the ratio $\dfrac{\text{change of velocity}}{\text{time taken}}$; so, in the limit,

$$f = \frac{dv}{dt} = \frac{d^2s}{dt^2}.$$

Force is the name given to the product of mass and acceleration, and

$$P = mf.$$

Fundamentally, none of these equations needs experimental proof; force *must* equal the product of mass and acceleration
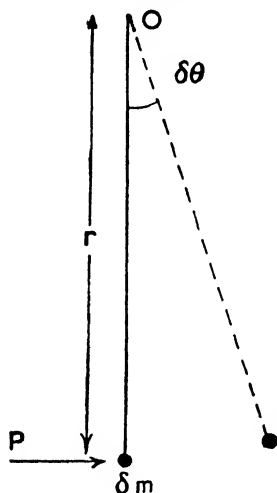
Fig 1

because force is merely the name given to the product of these two. It is, however, generally desirable to show that these scientific quantities are in accord with the vague 'popular' meaning of the term.

Most of the experiments 'to prove that force is equal to mass times acceleration' are really proving instead that the device for producing forces of differing strength is behaving accurately.

An equation which is precisely similar to $P = mf$ occurs also in angular mechanics. Consider a stiff weightless rod, pivoted at O, and carrying a point mass δm, as in Fig. 1.

The tangential acceleration $\dfrac{d^2x}{dt^2}$ produced by a force δP acting on the mass δm is given by

$$\delta P = \delta m \, \frac{d^2x}{dt^2}.$$

Now the product $\delta P \times r$ is called the couple acting on the system and so:

$$\delta C = \delta P \times r = \delta m r \left(\frac{r d^2\theta}{dt^2} \right)$$

since

$$x = r\theta$$

and

$$\frac{d^2x}{dt^2} = \frac{r d^2\theta}{dt^2}.$$

Now for a system of any shape $\dfrac{d^2\theta}{dt^2}$ will be the same for all parts of the body if this is solid, so

$$\int d\mathrm{C} = \left(\int r^2 \delta m\right)\dfrac{d^2\theta}{dt^2}$$

and

$$\text{couple} = \mathrm{I}\dfrac{d^2\theta}{dt^2}$$

providing I represents the total sum or integral of the products $r^2\delta m$ taken over the whole mass of solid object.

The fundamental equations of both straight line and angular mechanics are then both expressed in an identical mathematical form, so that the deduction of the remaining equations from these two primaries must be identical and may be summarized:

Straight Line Mechanics

$$\mathrm{P} = \dfrac{m d^2x}{dt^2} = m\dfrac{dv}{dt}$$

$$\therefore \underline{\int \mathrm{P}dt = \int m dv = mv_1 - mv_2}$$

Angular Mechanics

$$\mathrm{C} = \mathrm{I}\dfrac{d^2\theta}{dt^2} = \mathrm{I}\dfrac{d\omega}{dt}$$

$$\therefore \underline{\int \mathrm{C}dt = \int \mathrm{I}d\omega = \mathrm{I}\omega_1 - \mathrm{I}\omega_2}$$

Since, in collisions action and reaction are equal and opposite, at each instant $\int \mathrm{P}dt$ for one body must equal that for the other, and the momentum (product mv) lost by one body must equal that gained by the other.

$\therefore$ momentum must be conserved.

Now in angular collision experiments, in which, for example, one flywheel at rest is suddenly connected to another which is in motion, the couple produced at any instant by the one must equal that produced in it by the other.

$\therefore$ $\int \mathrm{I}d\omega$ before the collision must equal $\int \mathrm{I}d\omega$ after the collision and angular momentum must be conserved.

Kinetic Energy = work done	Rotational Energy = work done

$$= P dx$$

$$\therefore dE = P dx = m\frac{dv}{dt}dx$$

$$= mv\,dv$$

$$\therefore E = \tfrac{1}{2}mv_1^2 - \tfrac{1}{2}mv_2^2.$$

$$= P dx = (\text{couple})d\theta$$

$$\therefore dE = C d\theta = I\frac{d\omega}{dt}\delta\theta$$

$$= I\omega\,d\omega$$

$$\therefore E = \tfrac{1}{2}I\omega_1^2 - \tfrac{1}{2}I\omega_2^2.$$

$$\tfrac{1}{2}mv_1^2 - \tfrac{1}{2}mv_2^2 = E = \int P dx$$

$$\therefore v_1^2 - v_2^2 = \frac{2}{m}\int P dx$$

$$= \frac{2}{m}\int mf\,dx$$

$$\therefore v_1^2 - v_2^2 = 2\int f\,dx.$$

$$\tfrac{1}{2}I\omega_1^2 - \tfrac{1}{2}I\omega_2^2 = E = \int C d\theta$$

$$\therefore \omega_1^2 - \omega_2^2 = \frac{2}{I}\int C d\theta$$

$$= \frac{2}{I}\int I\frac{d\omega}{dt}d\theta.$$

$$\therefore \omega_1^2 - \omega_2^2 = 2\int \frac{d\omega}{dt}d\theta.$$

IF f IS CONSTANT

$$v_1^2 - v_2^2 = 2\int f\,dx$$

and $\underline{v_1^2 - v_2^2 = 2fs.}$

IF $\frac{d\omega}{dt}$ IS CONSTANT

$$\omega_1^2 - \omega_2^2 = 2\int\frac{d\omega}{dt}d\theta$$

$$\omega_1^2 - \omega_2^2 = 2\left(\frac{d\omega}{dt}\right)(\theta_1 - \theta_2)$$

$$\frac{dv}{dt} = \text{linear acceleration} = f$$

$$\therefore v_1 - v_2 = ft.$$

$$\therefore \frac{ds}{dt} = v_2 + ft.$$

$$\therefore s = v_2 t + \tfrac{1}{2}ft^2.$$

$$\frac{d\omega}{dt} = \text{angular acceleration} = \ddot{\theta}$$

$$\therefore \omega_1 - \omega_2 = \ddot{\theta}t.$$

$$\therefore \frac{d\theta}{dt} = \omega_2 + \ddot{\theta}t.$$

$$\therefore \theta = \omega_2 t + \tfrac{1}{2}\ddot{\theta}t^2.$$

§ 2. Simple Harmonic Motion

The subject of simple harmonic motion can be approached in three ways, viz. (1) by considering the projection on the diameter of a particle moving at a uniform speed on the circumference of a circle; (2) by regarding it as the motion of a body moving in a straight line towards a fixed point with an attraction proportional to its distance from the point; or (3) by defining it as a motion which can be represented by the equation

$$x = x_0 \cos pt.$$

In view of the mathematical relations existing between these definitions it would be quite immaterial which were adopted as the fundamental if it were not for the fact that it is much easier to differentiate than to integrate.

The functions $\cos (\omega t)$ and $\sin (\omega t)$ occur in the theory of the vibrations of stretched strings and it is from the connection of such strings
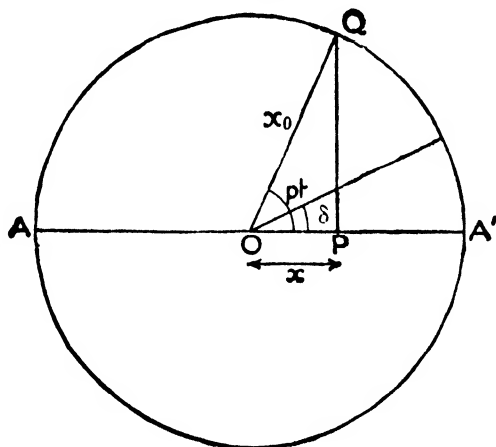
FIG. 2

with the musical scale that the use of the adjective *harmonic* has been extended to the motion of a point whose displacement is proportional to $\cos \omega t$ or $\sin \omega t$.

A simple harmonic motion was defined by Thomson and Tait in the words: 'When a point moves uniformly in a circle, the perpendicular QP, drawn from its position at any instant to a fixed diameter AA' of the circle, intersects the diameter

in a point P whose position changes by a simple harmonic motion.'

If this is accepted as our primary definition of the term Simple Harmonic Motion it follows *by deduction* that the equation

$$x = x_0 \cos (pt + \delta)$$

represents a simple harmonic motion (see Fig. 2).

From this equation *we further deduce* that the velocity of the point P is

$$v = \frac{dx}{dt} = -x_0 p \sin (pt + \delta),$$

and the acceleration is

$$\text{acc.} = \frac{dv}{dt} = \frac{d^2x}{dt^2} = -x_0 p^2 \cos (pt + \delta) = -x p^2.$$

The confusion regarding the definition of the term Simple Harmonic Motion has been largely produced by the many mathematical books which have a chapter heading 'Simple Harmonic Motion,' which ask the student to integrate the equation

$$\ddot{x} = -p^2 x,$$

which state that the result is a simple harmonic motion, and which finally *suggest* that simple harmonic motion was defined originally by the acceleration equation.

In consequence, students frequently waste much time and labour in purely physical discussions by proving the relationship between the two expressions in the wrong order, i.e. by integrating instead of differentiating.

Moreover, it must be fundamentally more correct to observe motion experimentally, then to obtain its equation, and finally to deduce the law of its force. For example, it is not generally possible to deduce from purely statical experiments on the forces produced by the extension of a spring whether any frictional forces will be called into play when the system is moving. Finally, it is worth mentioning that in atomic problems in which the particles are moving with speeds

comparable with that of light, the mass of a body increases with velocity; so that the motion would not be truly simple harmonic, even though the restoring force were accurately proportional to the displacement of the particle.

§ 3. Centrifugal Force

On page 6 we proved that the acceleration of the point P (Fig. 2) is towards the centre and is p^2x. The maximum acceleration of P is when it is on the circumference of the circle and is

$$p^2x_0.$$

This must also represent the acceleration towards the centre of the rotating point Q, since Q and P are then coincident.

We have, therefore, proved that the acceleration towards the centre of a particle rotating in a circle is p^2x_0, which is

$$\omega^2r$$

in the usual symbols.

§ 4. The Wave Equation

The equation

$$y = y_0 \cos p\left(t - \frac{x}{v}\right)$$

is employed throughout the whole of physics to represent the simplest type of wave motion. Although students are generally quite willing to employ the equation, they are generally distinctly uncertain why it really does represent a wave. We will consider this question under the following sections:

(a) Plot a graph of y against x keeping the time at a constant value t_a. The equation is then

$$y = y_0 \cos p\left(t_a - \frac{x}{v}\right).$$

The graph thus obtained looks the same as that of an ordinary water wave at any instant, and is shown in Fig. 3 by the full line.

(b) The wave moves. We prove this by drawing another graph corresponding to the time t_b. The equation is now

$$y = y_0 \cos p\left(t_b - \frac{x}{v}\right),$$

and the graph—shown by the dotted line—will go up and down in exactly the same way as the previous one.
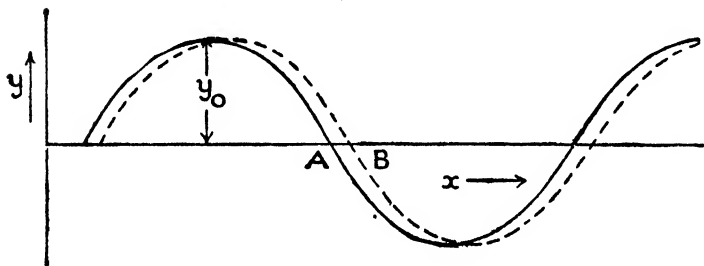
FIG. 3

Here, however, the value of x for which the displacement, y, is zero has changed from x_a to x_b.

Now since the displacement, y, and the values of $\cos p\left(t - \frac{x}{v}\right)$ are zero in the two examples, it follows that

$$p\left(t_a - \frac{x_a}{v}\right) = p\left(t_b - \frac{x_b}{v}\right),$$

and so

$$\frac{x_a - x_b}{v} = t_a - t_b.$$

Thus the wave has moved a distance $x_a - x_b$ in the time $t_a - t_b$, and must, therefore, have a velocity

$$= \frac{x_a - x_b}{t_a - t_b} = v.$$

We can now identify v as being the velocity of the wave.

(c) Next consider the way in which the displacement, y, at any particular value of x varies with the time. The equation of the motion, since x is constant, must be

$$y = y_0 \cos (pt + k),$$

and so the motion must be simple harmonic.

We see, therefore, that the graph of the equation looks like a wave, moves like a wave, and must be a simple wave since a point in the wave moves in a simple harmonic motion.

More complicated waves than this occur commonly in nature, and can be represented by more complicated functions of $\left(t - \dfrac{x}{v}\right)$, but the above is the simplest type of mathematical equation that can ever represent a wave. We therefore *define a simple harmonic wave as one which can be represented by the formula*:

$$y = y_0 \cos p\left(t - \frac{x}{v}\right),$$

in which y_0 represents the amplitude or maximum displacement, p the angular frequency, and v the velocity of the wave.

The displacement, y, will be at a maximum whenever $p\left(t - \dfrac{x}{v}\right)$ is equal to any integral number times 2π. Thus at any point x the displacement maxima will occur at intervals of T given by

$$pT = 2\pi.$$

Fourier's Theorem. The great importance of a simple harmonic wave lies in its convenient mathematical form. Perfect simple harmonic vibrations can never occur in nature because all natural vibrations both begin and end, while there is no term in the mathematical expression for a pure sine wave to lead one to expect that the wave should ever start or stop.

This difficulty is, however, overcome by regarding normal waves as collections of perfect simple harmonics with frequencies which differ very slightly, and which, therefore, combine together to produce a wave motion which both starts

and stops. As an illustration consider a wave of frequency f beating with another of frequency $f + n$. The beats must occur n times a second, since successive minima will occur whenever the two are exactly out of phase. This combination thus represents an approximately simple harmonic wave which starts and stops n times a second and which contains $\dfrac{f}{n}$ waves in each group (see Fig. 4).
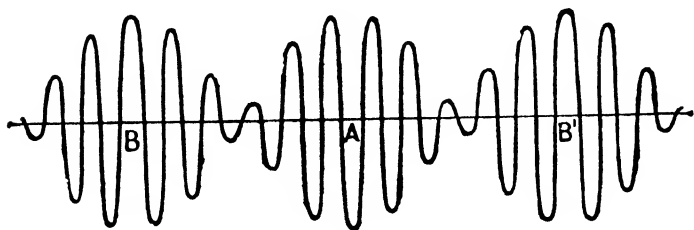
FIG. 4

By further combining this vibration with others of frequency $f + \dfrac{n}{2}, f + \dfrac{n}{4}$, etc., it seems possible that the wave systems B and B', C and C', D and D', etc., could all be made to vanish leaving only the simple wave group A.

Thus it appears probable that non-repeated wave trains having finite beginnings and endings can be regarded as combinations of suitable intensities of an infinite number of pure simple harmonic motions, whose frequencies differ from the fundamental by small amounts which should be of the order of

$$\frac{V}{\text{total length of wave train}},$$

since the length of the wave train $\sim \dfrac{f}{n}\lambda$.

A further problem now presents itself to our notice, since most waves which occur in nature have, in addition to a beginning and an ending, a very distorted wave form. It is, however, well known that even the most distorted wave form

can be produced by the combinations of suitable harmonic components of frequencies f, $2f$, $3f$, etc., providing that the intensities and phase relations between these components are correctly adjusted. Thus a normal wave of frequency f should be regarded as a complex mixture of waves of frequency $f \pm df$, $2(f \pm df)$, $3(f \pm df)$, etc., where the infinite collections of waves summarized by the term $\pm df$ take into account the finite length of the wave train.

We may now consider the limiting case of a single pulse vibration which can be regarded as a wave train of very short length and which should therefore be capable of synthesis from an infinite variety of simple harmonics chosen with suitable intensities and phase relations with any frequency whatever that may be required.

The general discussion given above illustrates a very important theorem proved accurately by Fourier. He showed that any periodic disturbance can be represented as the sum of sine waves of frequencies which are an integral times the fundamental, i.e.:

$$f \left\{ p \left(t - \frac{x}{v} \right) \right\} = \Sigma A_n \cos np \left(t - \frac{x}{v} \right),$$

in which A_n may be zero for some values of n.

He also showed that any disturbance, not periodic, may be represented as the sum of sine waves of *any* frequency.

Several interesting results follow from this; thus, we shall never be able to distinguish between a pulse and its equivalent collection of pure sine waves, once this pulse has been produced. Thus both the pulse and the complex wave mixture view of the nature of white light are really one and the same, and must both be equally correct.

§ 5. Stationary Waves. Mathematical Treatment

If two waves are passing simultaneously through the same space, the resultant displacement is the sum of the displacements contributed by each separately.

If the two waves have the same frequency and amplitude,

and if they are travelling in opposite directions, they can be represented by the following equations:

$$y_1 = y_0 \cos p\left(t - \frac{x}{v}\right) \rightarrow$$

$$y_2 = y_0 \cos p\left(t + \frac{x}{v}\right) \leftarrow$$

Thus the total displacement:

$$y = y_1 + y_2 = y_0 \cos p\left(t - \frac{x}{v}\right) + y_0 \cos p\left(t + \frac{x}{v}\right)$$

$$= \left(y_0 \cos pt \cos p\frac{x}{v} + y_0 \sin pt \sin p\frac{x}{v}\right)$$

$$+ \left(y_0 \cos pt \cos p\frac{x}{v} - y_0 \sin pt \sin p\frac{x}{v}\right)$$

$$= 2y_0 \cos p\frac{x}{v} \cos pt.$$

We thus see that a point in the medium moves with simple harmonic motion owing to the terms $\cos pt$. The disturbance can no longer move forward like an ordinary wave because it has no longer a term of the character $\left(t - \frac{x}{v}\right)$.
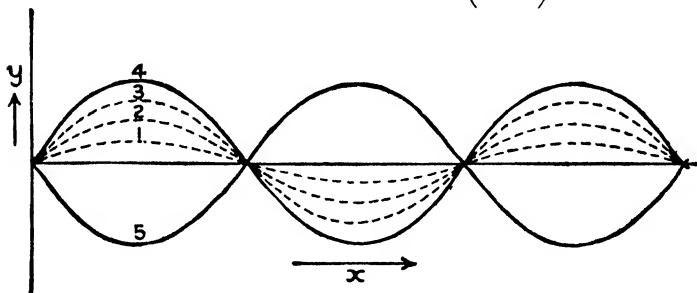
FIG. 5

In addition we see that the maximum value of the displacement, i.e. the amplitude, is $2y_0 \cos p\frac{x}{v}$, which varies from the maximum to zero and back to a maximum again as we

increase the distance x from the origin. Fig. 5 represents possible displacement values at slightly increasing times, $t = 0$, 1, 2, 3, or 4, etc., and shows also the limiting curves giving the maximum possible amplitude values.

§ 6. Stationary Waves.   Physical Treatment

In Fig. 6 (a), two waves which are moving in opposite directions and are of equal amplitude are represented by the full and the broken lines respectively; * the resultant displacement is shown by the dotted line as is indicated.
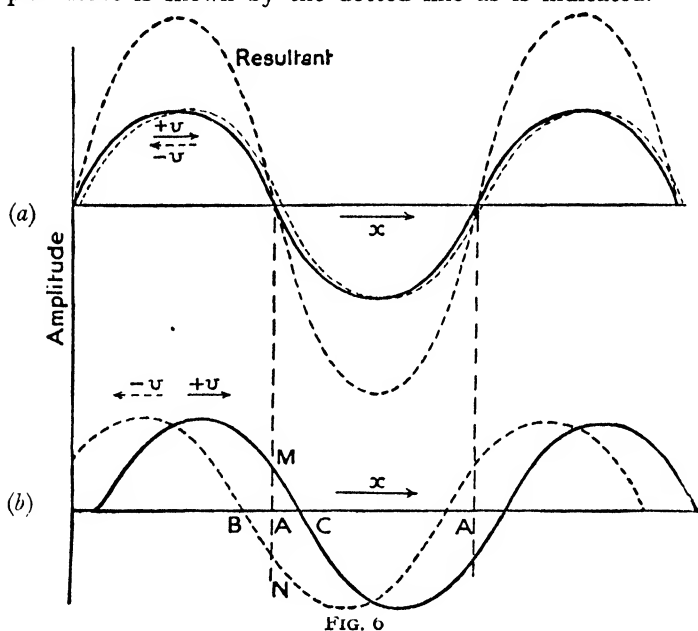
FIG. 6

Fig. 6 (b) gives the displacement values an instant later and shows that the waves have moved from A to B and from A to C respectively.

* For clarity the dotted curve has been moved slightly from the full curve in this Fig. 6 (a).   It should really be exactly superimposed on the full one at this instant.

Now if the velocities of the two waves are numerically the same, AB must equal AC and (if the amplitudes are also equal) AM must equal AN. Thus the combination of the two waves will still produce zero amplitude at the points A, A', etc.

This zero of amplitude at the point A will be quite permanent because the two lengths AM and AN will always be equal, since AB must always equal AC if the waves have the same velocities.

§ 7. Errors

If a very large number of experimental determinations are made of any one single quantity (such as the radius of curvature of a concave mirror by the 'one pin and parallax' method), the values obtained will not be absolutely identical, but will distribute themselves with equal numbers greater and less than the true value. A graph connecting the number of observations with the value obtained
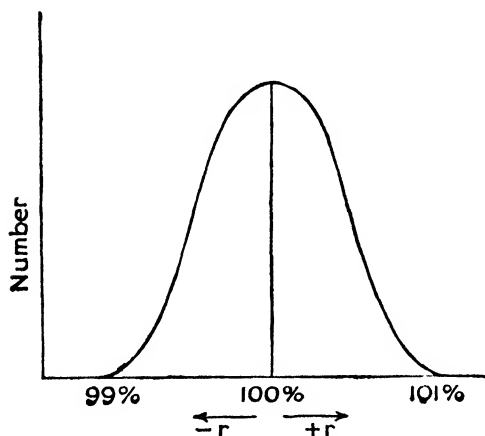
Fig. 7

has generally a form similar to Fig. 7 and can frequently be represented by the distribution law obtained by mathematicians for pure chance occurrences,

$$\text{probability} = \frac{1}{\sqrt{\pi}} a e^{-\frac{r^2}{a^2}},$$

where a is a number which is greater the more inaccurate

are the experiments, and where r represents the chance error.

In such practical experiments the results are frequently affected by systematic or constant errors so that the probability curve will be distributed about an incorrect value. In addition, it frequently happens that the probability curve is badly distorted through, for example, positive errors being more likely than negative ones.

It is, however, justifiable to employ formulae deduced from the ideal error formula to give approximate ideas of the probable error of practical experiments which are unaffected by systematic displacements.

The probable error is the error which is as likely as not to be exceeded, so that in a large number of values one half will have errors greater and one half less than this.

In addition, errors may be expressed in terms of 'mean errors' or 'root mean square errors,' whose meaning may be made clear by the following example, in which the true value is 100.

<div align="center">TABLE I</div>

Experimental Value								Mean
	104	97	100	99	98	103	99	100
Residuals	+4	−3	0	−1	−2	+3	−1	2·0
(Residual)^2	16	9	0	1	4	9	1	5·7

The *mean error* of each of the above values is the average value of the 'residuals' neglecting the sign, and is 2·0 in the above example.

The *mean square error* of each of the above values is the average value of the (residual)^2, and is 5·7 in the above example. The *root mean square error* is the square root of this mean square error, and is $\sqrt{5·7} = 2·4$ in the above example.

It is, on the whole, most common for experimentalists to calculate root mean square errors, but there is little advantage in performing this extra labour because mathematicians have

calculated the relation between these three expressions for errors.

Probable error = 0·674 (root mean square error)
= 0·845 (mean error).*

In practice, these three errors are completely mixed, so that some experimentalists calculate probable errors, some root mean square errors, and some mean errors, but almost all *call* the result the probable error.

§ 8. The Combination of Errors

Mathematicians state that the probable error in the mean of n independent values of a quantity is smaller than the error of each separate value by the ratio $\dfrac{1}{\sqrt{n}}$.

Thus in the above example the mean error of each value is 2·0; there are 7 separate values, so that the mean error of the average of these 7 is $\dfrac{2\cdot0}{\sqrt{7}}$.

This shows the advantage of taking a number of *independent* readings of any quantity. Four independent readings will give a mean value with an accuracy which is twice as great as that of one single observation. Nine readings will increase the accuracy by three, and sixteen readings by four.

Most physical quantities are obtained by multiplying or dividing two or more other quantities. In such cases the total error in the result is the sum of the *percentage* errors of each of the other quantities involved.

As an example consider the determination of g by the simple pendulum using the equation:

$$T = 2\pi \sqrt{\dfrac{l}{g}} \quad \text{or} \quad g = 4\pi^2 \dfrac{l}{T^2}.$$

* The probable errors calculated in this manner for the above example are 1·61 and 1·69 respectively. They do not agree—although they are derived from the same data—because the mathematical relationships are statistical and are only true if an infinite number of results are taken.

If the values *observed* were $l \pm dl$ and $T \pm dT$, the value of g obtained would be:

$$= 4\pi^2 \frac{l \pm dl}{(T \pm dT)^2} = 4\pi^2 \frac{l}{T^2} \frac{\left(1 \pm \dfrac{dl}{l}\right)}{\left(1 \pm \dfrac{dT}{T}\right)^2}$$

$$= \text{(true value of } g) \left(1 \pm \frac{dl}{l}\right)\left(1 \mp 2\frac{dT}{T}\right)$$

$$= \text{(true value of } g) \left(1 \pm \frac{dl}{l} \mp 2\frac{dT}{T}\right).$$

This example also shows that we must *add* the percentage errors even when the quantities are divided, because we do not know if the error is positive or negative.

§ 9. Independent Measurements

We can increase the accuracy of a single measurement by the factor $\sqrt{n}$ by taking n separate determinations only if these are completely independent. We may illustrate the meaning of the term 'independent' by another example.

If we were measuring the diameter of a penny by means of a simple vernier it would not be sufficient merely to adjust the vernier afresh for each determination, because we would *remember* the result of the previous measurement. We must not turn the penny round to other diameters because that would only tell us if the penny were truly circular, and we could not obtain independent readings by employing fresh vernier gauges, as this would merely tell us how uniformly the instruments had been made.

Instead it would be necessary to take a reading with a metal strip inserted between the penny and one jaw of the gauge, and to take another reading with the metal strip alone. The difference between the two readings would then give the diameter of the penny, and if each value had been obtained with strips of different thickness the results would be completely independent providing we did not perform the arithmetic in our head as we carried out the measurements.

B

CHAPTER II

GYROSCOPES

§ 10. The Vector Representation of Angular Velocities. § 11. The Representation of Angular Momenta. § 12. Precession. § 13. A Practical Gyroscope. § 14. Applications of the Gyroscope.

§ 10. The Vector Representation of Angular Velocities

In straight line mechanics it is possible to represent velocities vectorially by lines drawn in the direction of the velocity with

ω = 2 Radians per sec. ω = 1 Radian per sec.
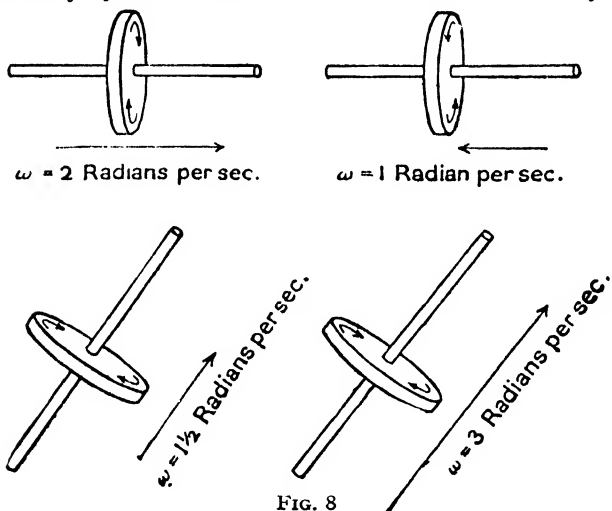
Fig. 8

lengths which are proportional to the magnitude of that velocity. In a similar way it is possible to represent by lines the three essential properties of angular velocities. These three essential properties:

(1) The body rotates about a definite axis;
(2) It rotates in a definite direction about that axis;
(3) It is of definite magnitude;

are represented (1) by the direction of the line, (2) by the direction of the arrows on the line, and (3) by the length of the line on any agreed scale. On the usual convention the rotation is that of a left-handed thread which is being screwed in the direction of the arrows on the line of the vector representation.

As examples of this method of representation, consider Fig. 8.

We will now consider an important theorem dealing with the composition of angular velocities and angular momenta.

Fig. 9 (a) represents an instantaneous picture of a flywheel and shaft which is mounted in a weightless ring and which is rotating in this ring about the axis AB with an angular velocity ω_1. In addition the ring itself is rotating about its axis CD with an angular velocity ω_2, so that if we had taken our photograph an instant later the point A would have been beneath the plane of the paper and the point B would. have been above it. We then say that the total angular velocity is compounded of the two component angular velocities ω_1 and ω_2 about the two axes AB and CD respectively.
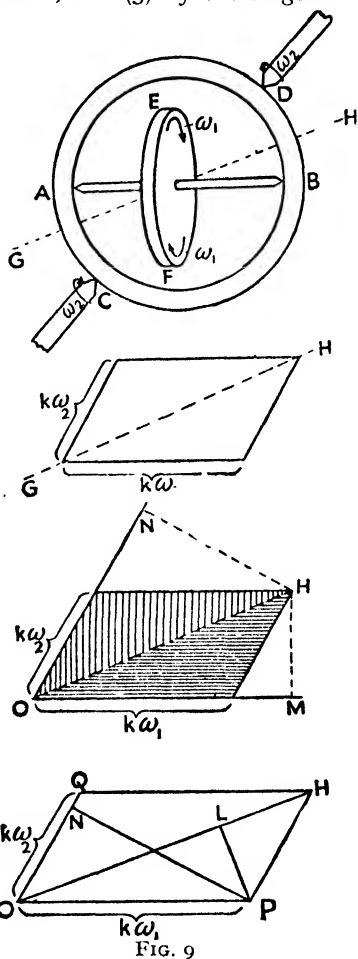
FIG. 9

Now at the instant that we took our photograph a part E of the flywheel (Fig. 9 (a)) will be moving into the plane of the paper and a part F moving out of the plane, so that, at that instant, the flywheel is effectively rotating about some other axis, such as, for example, GH. (Naturally this effective axis of rotation will have shifted an instant later so that G will then be beneath the paper and H above it.)

We will now prove that if the two angular velocities are represented vectorially, as in Fig. 9 (b), and if the parallelogram is completed as shown, then *the direction of this effective axis of rotation GH at the instant considered is that of the diagonal of the parallelogram OH.*

If this theorem is correct, the point H (and therefore the axis OH, since O is a fixed point) must be stationary at the instant we are considering. This means that if we refer to Fig. 9 (c):

$$\omega_1(\text{HM}) \text{ should equal } \omega_2(\text{NH}),$$

which must be the case since the two shaded triangles are equal.

Finally, we may prove that the line OH also represents— on our convention—the *actual magnitude* ω of the angular velocity of the flywheel about the axis OH at the instant we are considering. Thus, if we refer to Fig. 9 (d), the linear velocity of a particle at P is equal to ωPL and must also be ω_2PN (since P is on the axis of rotation of AOB). Thus:

$$\omega \text{ (PL) must equal } \omega_2\text{PN}$$

$$= \frac{1}{k} \left(\text{OQ.PN} \right)$$

$$= \frac{2}{k} \triangle \left(\text{OQH} \right)$$

$$= \frac{2}{k} \triangle \left(\text{OPH} \right)$$

$$= \frac{1}{k} \left(\text{OH.PL} \right)$$

hence $k\omega = \text{OH}$.

i.e. the length OH represents the angular velocity of the

flywheel at the instant considered, in direction and also in magnitude.

§ 11. The Representation of Angular Momenta

Since angular momenta have the same essential properties of direction and magnitude as have angular velocities, it follows that angular momenta compound by the same parallelogram law as angular velocities.

In the example of the rotating flywheel mounted in the rotating ring which has been discussed above, the resultant

Fig. 10.
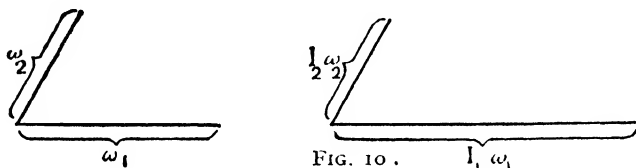
angular momentum obtained by the parallelogram law refers only to the flywheel and is distinct from the momentum of the rotating ring.

Moreover, since the angular momentum I_1 of a disc about a normal through its centre is greater than its value I_2 about a diameter, it follows that the angles involved in the parallelogram of angular velocities differ from those in the parallelogram of angular momenta. This is shown in Fig. 10.
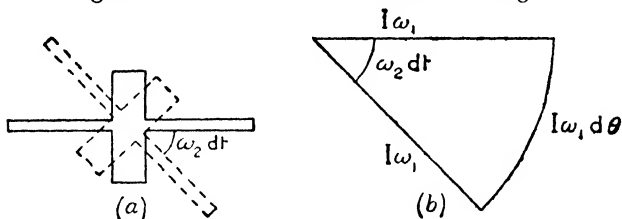
Fig. 11

§ 12. Precession

Let us suppose that the axis of a flywheel is rotating at a constant angular speed ω_2 so that after a time δt the axis has turned through an angle $\omega_2 \delta t$, as is shown in Fig. 11 (a). This

rotation of the axis of a gyroscope is described as a 'precession,' so that the gyroscope is said to be *precessing* with the angular velocity ω_2. (The same terms are also employed whatever may be the angle made between the axis of precession and the axle of the gyroscope flywheel.)

If the flywheel had been stationary instead of spinning, no couples would have been required to maintain this movement of the flywheel shaft because it would have been turning with a constant speed and neither accelerating nor retarding.

A reference to the vector representation of Fig. 11 (*b*) shows, however, that the direction of the angular momentum of the flywheel has changed, so that an angular momentum $I\omega_1\delta\theta$ or $I\omega_1\omega_2\delta t$ must have been supplied to the system in the time δt. Thus, since $I\delta\omega = C\delta t$, a couple C must have been applied such that

$$C\delta t = I\omega_1\omega_2\delta t$$
$$\text{i.e.} \quad C = I\omega_1\omega_2.$$

The whole theory of gyroscopes and of the practical applica-
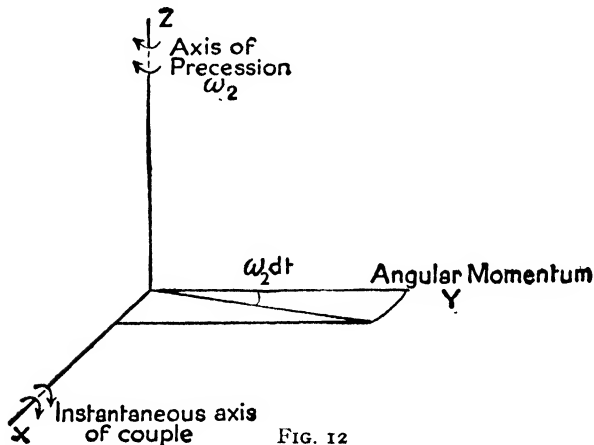
FIG. 12

tions turns on this theory and on the fact that *this couple must be applied at right angles to the plane of the paper so that its axis is in the plane of the paper.*

We have, therefore, shown that, to turn the flywheel axis *at a constant speed* in the direction shown, no couple is required in the plane of the paper if the flywheel is stationary, or rotating, whereas if the flywheel is rotating a couple of magnitude $I\omega_1\omega_2$ is required at right angles to the plane of the paper.

These relations are exhibited in Fig. 12.

§ 13. A Practical Gyroscope

A practical gyroscope might consist of a flywheel mounted in a ring which is pivoted at the point A to a vertical rod (Fig. 13). The pivot allows the flywheel and ring to move vertically, and, when the screw at B is free, the system can also rotate about the vertical axis AB.

In the position shown, gravity is producing a couple, $Mgl \sin \theta$, so that if the gyroscope can once be made to precess about the vertical axis AB with an angular velocity ω_2, this precession will continue uniformly, providing:

$$I\omega_1\omega_2 = C = Mgl \sin \theta.$$

A uniform precession of this nature will only occur if the gyroscope is made to precess with this angular velocity ω_2 at the exact instant that it is released and the full gravitational couple is applied. If, however, the gyroscope is released without being given this precessional motion, the axis will begin to drop, a couple will be produced tending to rotate the centre of gravity about the axis AB, and the system will commence to precess with a rapidly increasing speed until the gravitational couple is exactly equal to $I\omega_1\omega_2$.

In cases of this kind oscillations frequently occur with the
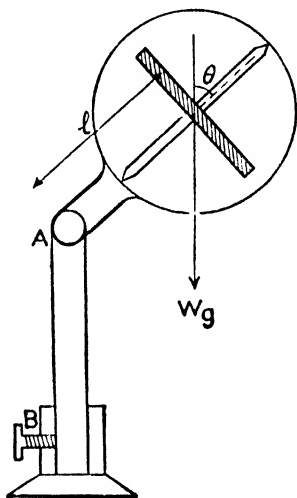
FIG. 13

flywheel axle alternately sinking below and rising above its equilibrium position.*

This example proves, moreover, the general theorem that *when a couple is applied to try to rotate the axis of a gyroscope or rotating flywheel, the axis moves instead at right angles to the expected direction and tries to set in the same direction as, and to move in the same sense as, the axis of the applied couple.*

One further general property of gyroscopes can be illustrated by this example. The principle of the conservation of energy shows that when the gyroscope was released with zero precessional energy its centre of gravity must drop by a small distance to supply the final energy of the precessional motion $\frac{1}{2}I_2\omega_2{}^2$, since the nature of the flywheel bearings prevents the possibility of the required energy being taken from the flywheel. Moreover the principle of conservation of angular momentum shows that the total sum of the precessional angular momentum and the *component* of the momentum of the flywheel about the vertical must always be constant. Thus as the gyroscope dies and ω_1 becomes smaller, the precessional momentum and the precessional energy must increase. Once again the energy necessary for this increase in precessional velocity is obtained from the loss of height of the centre of gravity of the flywheel.

§ 14. Applications of the Gyroscope

(*a*) *Schlick's apparatus for the stabilization of ships.* The principle of this apparatus may be gathered from Fig. 14, which represents in a general way a transverse section of a ship and a gyroscope arrangement. When the ship rolls with the points A and B alternately rising above and falling below the horizontal, the gyroscope and ring will 'wobble' about the axis AB. If this axis is provided with suitable friction discs the energy of the ship's roll will be rapidly removed.

* Notice that the gyroscope can only withstand the couple if it is allowed to precess, and that if the precessional velocity ω_2 is zero the couple it can support will also be zero.

(b) *Brennan's Monorail.* Louis Brennan employed a gyroscope to keep upright a passenger-carrying car which ran on one single rail with two wheels like a tandem bicycle. In
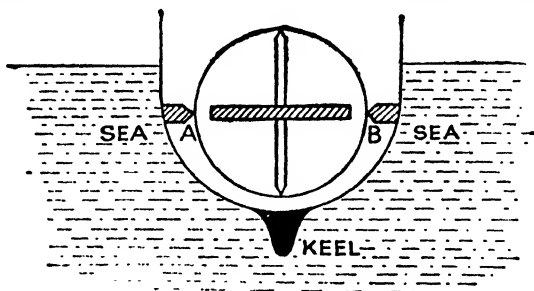
FIG. 14

this apparatus the gyroscope was mounted as shown in Fig. 15 with its axis parallel to the axles of the car wheels. If the car tilts slightly to one side the gyroscope will precess about the vertical axis, and in consequence, if the precession is *increased* artificially by the operator, the car will right itself.
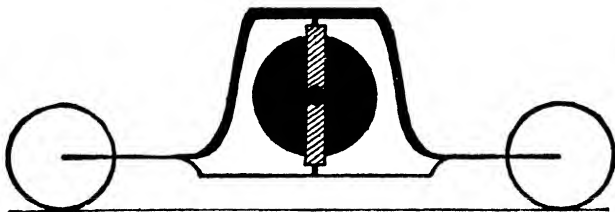
FIG. 15

An automatic method of increasing the precession whenever the system tilted was finally arranged, so that the car was held perfectly stable even when it was stationary.

(c) *The Gyroscopic Compass.* Magnetic compasses are made somewhat uncertain by the variations which occur in the earth's magnetism and by the irregular magnetization of the ship itself. Now a gyroscope which is mounted on gimbals, so that it can point freely in any direction, will remain with

its axis in a fixed direction in space. Such an instrument would not be of great utility, because if the flywheel ceased to rotate or encountered any resistance the direction of the axis would have to be determined afresh by astronomical observation. If, however, we arrange that the ring which supports the axle of the flywheel is horizontal, by placing the centre of support above the centre of gravity we get an interesting result.
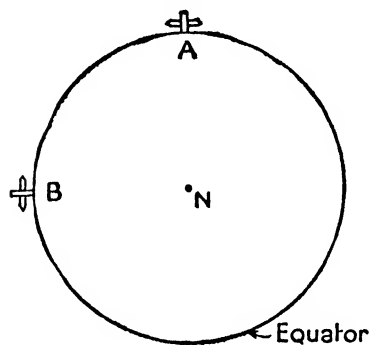
Fig. 16

Fig. 16 represents the earth *as seen from the pole star*. Suppose that at one instant the gyroscope is at the point A and has its axis pointing east and west as shown. Six hours afterwards it will be at the point B, and we should at first expect the axis to be turned through 90° as shown in the figure (i.e. still horizontal and still pointing east and west). We notice, however, that to turn this axle couples must have been supplied by the mercury bath in which the instrument is floating horizontally, and we then remember that when a couple is applied to alter the direction of the axis of a rotating gyroscope the latter immediately precesses so as to set with its axis of rotation pointing in the direction of the axis of the couple. In this example, therefore, it follows that the gyroscope will precess until it finally sets with its axis of rotation pointing north and south.

CHAPTER III

PROPERTIES OF MATTER

§ 15. Gravitation. The Relative Accuracy of the Cavendish and the Boys Experiment

Although most text-books on properties of matter include descriptions of the Cavendish experiment to find the gravitational constant of matter, and of Boys's modification of it, students are usually uncertain why the latter experiment is the more accurate. This single point, therefore, alone warrants consideration here in a manner which presumes that the student has first referred to the normal text-books, and has obtained the full descriptions of the experiments. As is well known, the first advantage of the Boys experiment lay in its small size which permitted more accurate temperature equilibrium to be reached, since this meant that convection currents were largely obviated. In addition the Boys experiment has fundamentally a greater sensitivity.

Fig. 17 is a simple plan of the apparatus employed in the Cavendish experiment and shows the symbols which are employed here. The force of attraction (in dynes) between a large sphere and its adjacent small one is $G\dfrac{Mm}{d^2}$, since

uniform, spheres act as if their mass is concentrated at the centre.

The couple produced by the two attractions is therefore $G\dfrac{Mm}{d^2} \times l$ dyne cm. This couple must be balanced by the
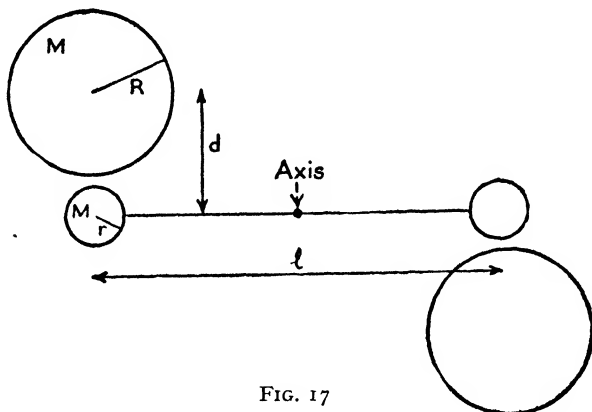
FIG. 17

couple produced by the rotation of the suspending fibre $(\mu\theta)$, and so

$$\mu\theta = G\frac{Mm}{d^2}l$$

and

$$\theta = G\frac{Mm}{d^2}\frac{l}{\mu}.$$

Although this expression for the angle of deflection is mathematically correct it is *most undesirable* for a physicist to leave it in this form, because it gives an entirely false impression of the factors on which the angle of deflection depends. This happens because it is not expressed completely in terms of quantities which are actually measured. Thus μ is always found by noting the time of oscillation (T) of the system and using the equation:

$$T = 2\pi \sqrt{\frac{I}{\mu}},$$

and so
$$\mu = \frac{4\pi^2}{T^2} I = \frac{4\pi^2}{T^2} \frac{(ml^2)}{(2)},$$

if we take as an approximation that $I = \frac{ml^2}{2}$. (This approximation is somewhat rough in the Boys experiment where the length l is very small.)

We have, therefore:

$$\theta = G\frac{Mm}{d^2} l \left(\frac{T^2}{4\pi^2} \frac{2}{ml^2} \right)$$

$$= G\frac{M}{d^2} \frac{T^2}{2\pi^2} \frac{1}{l}.$$

Thus we see that, when the formula is expressed in terms of quantities which are actually measured, the angle of deflection becomes greater as the length l becomes smaller.*

This complete change in the general impression of the most suitable values to employ in the experiment is an excellent example of the modern tendency to express a mathematical formula in terms of those quantities which are actually to be measured.†

* Several text-books consider a decrease in linear dimensions applied to every part of the Cavendish apparatus and reach the conclusion that the sensitivity is independent of the linear dimensions. Although this is quite correct, it does not apply in this instance since Boys did not reduce the radii and masses of his attracting balls in the same ratio as he did the lengths of the arms.

† Another good illustration of the false impressions which may be gathered when a formula is not expressed throughout in measured quantities is to be found in connection with Poiseuilles's method of determining the viscosity (η) of a liquid by observing the volume (V) flowing in a time t through a long capillary tube of length l, and radius r. The formula usually quoted is

$$\eta = \frac{p\pi r^4 t}{8lV}$$

where p is the pressure difference between the two ends of the tube. Students almost invariably regard this experiment as being essentially a determination of the rate of liquid flowing through the tube, whereas, when the radius of the tube is expressed in terms of the measured length and mass of mercury which fills it, the accuracy of the determination is found to be governed more by that of the length of the tube than by any other factor.

In order to obtain the maximum value of the deflection, both Cavendish and Boys arranged their apparatus to have as large a value of T as was practicable, i.e. 4–7 minutes.

In addition, both experimenters employed spheres of sizes which gave the largest practicable value to $\frac{M}{d^2}$, but were both limited by the same conditions that $M = \frac{4}{3}\pi R^3 \rho$ and d could not be smaller than $R + r$. Thus $\frac{M}{d^2}$ was very nearly equal to $\frac{\frac{4}{3}\pi R^3 \rho}{(R+r)^2}$, which is $\frac{4}{3}\pi R \rho$ approximately. The accuracy therefore increases proportionally to the radius and to the density of the spheres employed. Both experimenters used lead, which is the densest material obtainable in bulk (except for mercury). Cavendish used spheres of 6 in. radius which weighed 370 lb., he could therefore have increased his value of $\frac{M}{d^2}$ by a factor of 2 only by increasing R by 2, which would have involved increasing the weight to about 3,000 lb. of lead.

Boys used spheres of only $2\frac{1}{4}$ in. radius and 20 lb. weight; he, therefore, had an angular sensitivity which was smaller, in virtue of $\frac{M}{d^2}$, than if the larger spheres had been employed, by a factor of $\frac{6}{2\frac{1}{4}}$. In actual practice Boys had nearly the same value of $\frac{M}{d^2}$ as Cavendish; partly because he employed gold spheres of only 0·2 in. diameter for the suspended balls instead of the 2-in. lead spheres of Cavendish, and partly because he placed the spheres rather closer together than did Cavendish.

The outstanding reason for the superiority of Boys's experiment lies in the last term in the expression for θ, since Boys had a value of l of only 0·9 in. while Cavendish used 6 ft. Boys avoided cross-attractions between suspended masses by

the well-known scheme of arranging one of his swinging balls to move in a horizontal plane well below the other.

It is very possible that Cavendish employed his large value of l not because he expected it would give greater sensitivity—as many people incorrectly imagine—but because he was unable to obtain a fibre which would give such a long time of oscillation with a short rod. The difficulty is, that to obtain a large time of oscillation, T, the ratio $\dfrac{ml^2}{\mu}$ must be large and must also be subject to the condition that the μ cannot be made smaller than some definite limit by reducing the diameter of the torsion wire, without making the wire so fine that it is unable to support the weight of the small balls ($2m$).

The most suitable material for the fibre is, therefore, one which combines a small value of the rigidity with a very large tensile strength. Cavendish used silvered copper wire for which

$$\frac{\text{rigidity}}{\text{tensile strength}} = \frac{4 \cdot 5 \times 10^{11}}{1 \cdot 3 \times 10^{9}} = 340;$$

while Boys used quartz fibres with a ratio of

$$\frac{3 \cdot 0 \times 10^{11}}{10^{10}} = 30.$$

In addition, the maximum value of m which can be supported by any fibre depends on the square of the radius of the fibre, while the torsional constant of the wire μ depends on the fourth power of the radius. There is, therefore, an inherent advantage in using exceedingly fine fibres and correspondingly small values of m.

The superiority of the Boys experiment over that of Cavendish therefore depends essentially on the technical achievement of producing fibres of fused quartz of diameters down to 10^{-4} cm. with relatively low rigidity and high tensile strength.

It is of interest to notice some of the methods by which these quartz fibres may be produced. A piece of commercial fused quartz rod is pulled out in an oxyhydrogen flame to

the shape shown. Boys made some of his fibres from such pieces by attaching one end to an arrow and holding the other end stationary. A portion of the fine rod was then softened again in the oxyhydrogen flame and the arrow
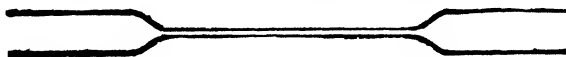
FIG. 18

immediately shot from a bow. The plastic quartz was then drawn out into a very long thread or fibre.

On a commercial scale the one end is fixed to a large drum connected to an electric motor, which is started at the instant that the quartz becomes plastic. Twenty to one hundred yards of quartz are thus wound on the drum before the quartz solidifies again.

An excellent method for the production of fibres of all sizes is to snap in two one of the pieces mentioned above, to melt a thread end in an oxyhydrogen flame, and to feed in more and more of the fine rod until a long tail of soft quartz is formed in the flame itself, as shown in Fig. 19.
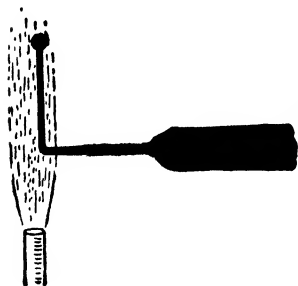
FIG. 19

If, now, the gas velocity of the oxygen and hydrogen of the jet is great enough, the small 'blob' at the end will suddenly be pulled off with the gas stream, will be shot across the room, and will carry a long fibre of quartz behind it. Owing to the difficulty in finding these threads it is generally best to erect a large screen of *black* velvet some distance in front of the jet and to produce ten or twenty fibres first before commencing the search. Very fine fibres may float in the air for a long while before coming to rest, and are so fine that they are quite invisible unless strongly illuminated on a perfect black surface.

§ 16. Gravitation.   Modern Work

Boys almost completely exhausted the accuracy possibilities of the torsion balance method for the determination of the gravitational constant of matter, G. This method

(a)   FIG. 20   (b)
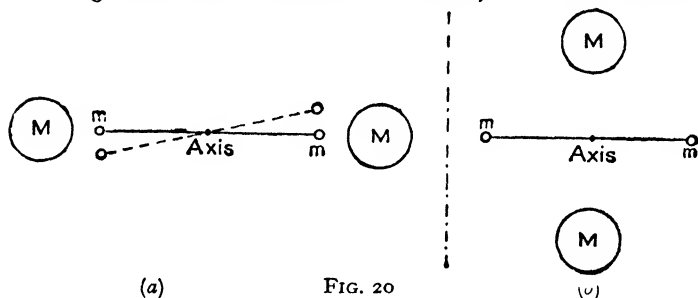
has now been completely superseded by experiments in which the time of swing of two spheres at the ends of a horizontal rod is measured—a method which was first introduced by Braun, and was extended by Heyl in 1927.

In Fig. 20 (a) the rod and small masses m will tend to oscillate about the central axis as shown under the gravitational
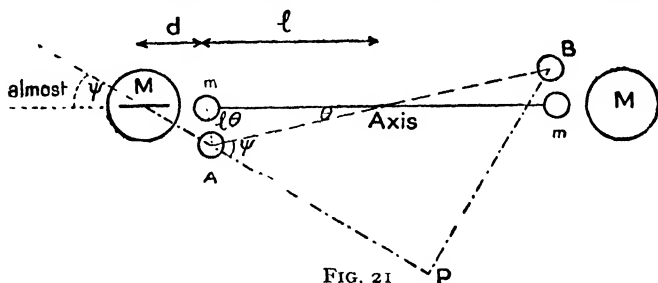
FIG. 21

attraction of the larger masses M and the normal rigidity of the suspending wire.   In Fig. 20 (b) the suspending wire alone produces the restoring force, so that by comparing the times of oscillation in the two experiments it should be possible to determine the gravitational attraction between the masses when in position (a) and thus to obtain G.

In Fig. 21 the attractive force between M and m is $G\dfrac{Mm}{d^2}$ dynes, and will be almost the same whether the small mass is in its zero position or in its displaced position A.

The couple produced by these two forces when in the displaced position will be:

$$\left(G\frac{Mm}{d^2}\right)(PB) = \left(G\frac{Mm}{d^2}\right)\left(2l\sin\psi\right)$$

$$= \left(G\frac{Mm}{d^2}\right)(2l)\left(\theta + \frac{l\theta}{d}\right) \text{ approximately.}$$

Thus, instead of our usual equation for the angular acceleration produced by a wire under torsion:

$$I\ddot{\theta} = -\mu\theta,$$

we have now:

$$I\ddot{\theta} = -\left\{\mu + \left(G\frac{Mm}{d^2}\right)(2l)\left(1 + \frac{l}{d}\right)\right\}\theta$$

Thus the new time of oscillation T is given by the expression:

$$T = 2\pi\sqrt{\frac{I}{\mu + G\dfrac{2mMl}{d^2}\left(1 + \dfrac{l}{d}\right)}},$$

from which we see that

$$\mu + G\frac{2mMl\left(1 + \dfrac{l}{d}\right)}{d^2} = \frac{I4\pi^2}{T^2}.$$

It follows that the time of oscillation T_0 when the gravitational attraction is absent (as in Fig. 20 (b)) is larger than T, and that

$$G\frac{2mMl\left(1 + \dfrac{l}{d}\right)}{d^2} = I4\pi^2\left(\frac{1}{T^2} - \frac{1}{T_0^2}\right).$$

In this expression the times, the attracting mass, and the distances can be measured experimentally, so that G can be determined.

In Heyl's experiments l was 10 cm., m 54 gm., M an iron cylinder of 66 kg. weight, 28·5 cm. long and 19·5 in. diameter. The suspending filament was one metre in length and was a tungsten wire 0·025 mm. diameter (i.e. the *radius* was ¼ and the area of cross-section $\frac{1}{16}$ of that of No. 40 copper, which is the finest wire normally encountered in the laboratory. Electric lamp wire—the wire, and not the coils we normally consider—varies in radius from 0·007 to 0·025 mm.).

Tungsten was chosen because it is very reproducible, is easily tied, and does not 'break for no apparent reason' in the same way as the fused quartz fibres that Boys employed.*

The times of swing obtained by Heyl were respectively 29 and 34½ *minutes*, and his final value of G was

$$G = 6·664 \pm 0·002 \times 10^{-8} \text{ dyne cm.}^2 g^{-2},$$

which is over five times as accurate as the value obtained from the Boys experiment.

The experiment was carried out in the constant-temperature room at the American Bureau of Standards and was 35 ft. below ground. Cylinders were employed owing to the experimental difficulty of making spheres, although this necessarily meant much difficulty in the calculation of the results. The cylinders were cast 12 in. in diameter and were forged down to 9 in., in order to close blow-holes and to make the metal of uniform density. For a similar reason the small masses m were obtained by fusion in a vacuum.

The apparatus was set swinging by employing the gravitational attraction of 2 kg. of mercury. For ¼ of a swing the bottles were placed on one side, and for the next ½ swing on the other side. It was possible in this way to build up a total amplitude of 3½° in 2 hours. The position of the oscillating beam was observed by reflecting light from a light mirror mounted on the beam.

* The tungsten wire used in making electric lamp filaments is very fine and very ductile, but it is quite impossible to employ electric lamp filaments as a source of ductile tungsten for the laboratory, because once the filament has been heated the metal crystallizes and becomes exceedingly fragile and brittle. This explains why partly used electric lamps need much more careful treatment than new ones.

It is of interest to notice that the theoretical expression shows that the value of $\left(\dfrac{1}{T^2} - \dfrac{1}{T_0^2}\right)$ should be nearly a constant for all experimental arrangements since M is proportional to the cube of its radius, and this radius will be very nearly identical with d if the large mass is placed close to the smaller ones. This difference will not normally be detectable with large scale torsion systems because the time of oscillation is too small. As an example we may note that if T = 60 sec. the corresponding difference in times would be only $\dfrac{6}{100}$ sec. It should, however, be quite possible to make models of the apparatus, using electrically charged spheres in which the attractive force $G\dfrac{mM}{d^2}$ is replaced by the similar expression $\dfrac{e_1 e_2}{d^2}$.

§ 17. Gravitation. Uniform Spheres gravitate as if their Mass were concentrated at the Centre

Many experiments and problems contain the implicit assumption that uniform spheres act as if their masses were concentrated at the centres. Reasons for this assumption may readily be deduced by relying on the arguments which apply in electrostatics without going into the details of Newton's formal proof.

The electrostatic law of force $\left(\text{force} = \dfrac{e_1 e_2}{kd^2}\right)$ is mathematically identical with the gravitational law of force except in regard to the sign of the attraction or repulsion. Any deductions based solely on the electrostatic law of force must therefore apply, with a change of sign, to the gravitational law of force.

Now in electrostatics we develop the conceptions of fields of force, of lines of force, of lines of induction, and of Gauss's theorem, and use this theorem to find the number of lines of induction crossing unit area of a mathematical sphere sur-

rounding and concentric with a uniformly charged sphere. We then show that the number of lines of induction is always the same as would have been obtained if the whole of the electrostatic charge had been concentrated at the centre.

In view of the exact mathematical analogy between the two laws of force we may employ this result directly and state that a thin spherical shell of uniform mass must also act as if its mass were concentrated at its centre. If desired, one could clearly develop for gravitation also the same conceptions of fields of gravitational force, lines of gravitational induction, and Gauss's gravitational theorem, and develop the proof without introducing the word 'electricity' at any stage.

Once we have proved that a uniform material spherical shell acts on external objects as if its mass were concentrated at the centre, it follows directly that any concentric series of such shells must also act as if their mass were concentrated at the centre.

It is not necessary for the shells to be all of the same density providing each individual shell is uniform, so that the earth acts as if its mass were concentrated at its centre, even though its core is certainly much denser than the material on the surface.

§ 18. Gravitation. The Earth attracts Objects inside it with a Force proportional to the Radius

In electrostatics we have the formal theorem that there is no electrostatic force inside an isolated *uniformly charged* sphere (or inside *any* closed surface of uniform *potential*).

In a similar way it follows that there is no gravitational force inside a uniform gravitational shell, or concentric series of shells. Thus the only force exerted on a particle inside the earth at a distance R from the centre is due to the sphere of radius R, and that part which is at distances greater than R can have no effect. Thus:

$$\text{acceleration on particle} = \frac{G(\frac{4}{3}\pi R^3 \rho)}{R^2} = \frac{4}{3}\pi G R.$$

§ 19. Gravitation. Tides

Since the moon produces *two* tides a day on the earth, the state of the tides at any instant can be represented as in Fig. 22.
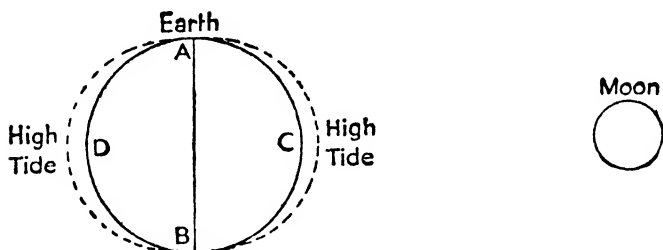
FIG. 22

If we disregard the very slight tides which occur in the solid material of the earth itself, we may say that the earth acts as if its mass were concentrated at its centre and that it ·is balanced between its gravitational acceleration towards the moon and its centrifugal force away from it.

The high tide at C is then caused by the water in the part ABC being rather nearer the moon than the centre of the earth, and being, therefore, attracted towards it with a slightly greater acceleration.

Correspondingly the high tide at D is caused by the water in the part ADB being further from the moon than the centre of the earth and so having a smaller gravitational acceleration. In consequence, the centrifugal force is the greater and the water is flung into a heap on the part of the earth furthest from the moon.

§ 20. Surface Tension. Condensation into Drops

It is proved in the usual text-books that the vapour pressure over a concave liquid surface of the type which occurs in a

wet capillary tube is less than that over a plane surface by an amount *

$$= \frac{2T}{r} \frac{\text{density of vapour}}{\text{density of liquid} - \text{density of vapour}}.$$

The vapour pressure of a spherical *drop* of liquid of radius r must be *greater* than that over a plane surface by the numerical amount given by this formula, because it is a *convex* surface and has a radius which is opposite in sign to that over a capillary meniscus.

In calculating numerical values we must remember that T is expressed in dynes per cm. length, so that δp will be in dynes per sq. cm., and not in cm. of mercury. It is interesting to notice that at 15°, when the vapour pressure over 'flat' water is 12·8 mm. of mercury, the vapour pressure is greater over a drop of radius r by $\dfrac{1\cdot2 \times 10^{-6}}{r}$ mm. of mercury. Tiny drops of radius 10^{-7} cm. will thus have a total vapour pressure which is twice as great as that over a plane surface. The whole effect will, therefore, be of insignificant importance with drops as large as rain or ordinary cloud or mist particles, will be of *interest* in connection with very small mist drops, but will be of *vital importance* in the birth of such drops from the chance encounters of several water-vapour molecules, since the radius will then be of the order of 10^{-7} cm.

* This theorem is proved by considering the equilibrium which must finally occur between liquid and vapour in the system shown in Fig. 23. The vapour pressure immedia y outside the meniscus must be equal to that at the point A, and so must be less than that at the point B (which is the vapour pressure over a *flat* liquid surface) by an amount equal to $h \times$ (density of vapour). Now h—the rise of the liquid in the capillary tube—may be calculated in the usual manner by the equation:

$$h \text{ (density of liquid} - \text{density of vapour)} = \frac{2T}{r};$$

so that the decrease in vapour pressure over the concave meniscus of radius r is

FIG. 23

$$\frac{2T}{r} \frac{\text{density of vapour}}{\text{(density of liquid} - \text{density of vapour)}}.$$
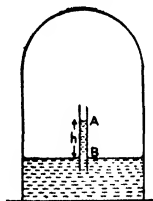
Since it must be exceedingly uncommon for any *large* number of molecules to arrive simultaneously at the same point and thus to form a drop of finite size as a result of the chance encounters, it follows that cloud or mist drops will only form readily on dust or similar particles with radii greater than 10^{-6} or 10^{-7} cm.

If the dust particles have radii of 10^{-7} cm., the above example shows that condensation will only occur at 15° C. if the actual water-vapour pressure is double its saturation value for a plane surface, while if the radius is 10^{-6} cm. the actual water-vapour pressure must be only 1·1 times as great.*

§ 21. Surface Tension.   Condensation on Dust Particles

This selective condensation of water on dust particles has been used in the measurement of the number of dust particles in the atmosphere. The difficulty of obtaining water vapour supersaturated by any desired amount is surmounted by expanding damp air adiabatically with a suitable cylinder and piston. The theory of the method is conveniently explained by an example.

The water-vapour pressure in saturated air at 15° C. is 12·8 mm., so that if the air is expanded to 1·4 times its previous value, the actual water-vapour pressure in the expanded air will be approximately $\dfrac{12\cdot8}{1\cdot4} = 9.2$ mm.   If this expansion of

---

* It is interesting to note that the 'bumping' which some liquids exhibit on boiling may be explained in the same manner. The equilibrium vapour pressure inside a small bubble of vapour which is surrounded by liquid will be *greater* than that over a plane surface by:

$$\frac{2T}{r} \left(\frac{\text{density of vapour}}{\text{density of liquid} - \text{density of vapour}} \right).$$

Thus if water is to form bubbles of size $r = 10^{-7}$ cm., the vapour pressure must be (76 + 54) cm. of mercury, so that the water must be heated to 116° C.

Once a bubble of this size has formed it will increase in size explosively, because the vapour pressure of the surrounding water will be that corresponding to 116° C., while its own equilibrium vapour pressure will drop as it gets bigger, and will be (76 + 5·4) cm. of mercury at $r = 10^{-6}$ cm., and (76 + 0·54) cm. of mercury at $r = 10^{-5}$ cm.

the air is performed adiabatically, the temperature will drop by an amount which can be calculated from the equations:

$$p_1v_1^\gamma = p_2v_2^\gamma$$
$$p_1v_1 = RT_1$$
$$p_2v_2 = RT_2$$

to be 252° A. or −21° C. At this new temperature, the saturation vapour pressure of 'plane water' is 0·78 mm. Thus the air actually contains 9·2 mm. pressure of water, whereas it should only hold at this temperature 0·78 mm.*

In the actual experiment of determining the total number of dust particles per cc. of laboratory air, it is usual to obtain dust-free air first by carrying out a large number of such expansions and allowing time for the water drops to sink (and adhere) to the floor of the expansion chamber. A small quantity (0·1 cc.) of the dusty laboratory air is then diluted with a larger quantity (10 to 100 cc.) of dust-free air, and 0·1 cc. of this mixture admitted into a known volume (500 cc.) of air in the expansion chamber. An expansion is then made, the supersaturated water vapour condenses and forms drops on the dust particles, the drops are strongly illuminated, and are immediately photographed. The number of dust particles in the air is then obtained by direct counts of the drops recorded on the photographic plate.

The somewhat complicated procedure of the preliminary dilution of the air is generally necessary in view of the very large number of dust particles (10^5 to 10^6 per cc.) normally present in the air.

§ 22. Surface Tension. Condensation on Ions

C. T. R. Wilson has shown that condensation into drops can occur on gaseous ions if the supersaturation exceeds a certain value.

* C. T. R. Wilson has shown that if the supersaturation is *sufficiently great*, dense clouds can be formed in the absence of dust particles. This occurs, however, only if the expansion is greater than the value 1·4 times employed in the above example. For this value of the supersaturation (12 times) the critical radius is of the order of the size of a water molecule.

This is presumably to be correlated with the repulsion which exists between charges of the same sign. The forces due to surface tension tend to make a drop become smaller and diminish its surface area, while the forces due to electrostatic repulsions tend to enlarge the sphere and to increase its surface area. Thus electric charges must reduce the surface energy and must lower the effective surface tension.

The mechanical work done against the forces of surface tension in forming a sphere of radius r is:

$$4\pi r^2 T.$$

The electrostatic energy of a charged drop is $\dfrac{e^2}{r}$, so that the total energy in virtue of these two effects is:

$$4\pi r^2 T - \frac{e^2}{r}$$
$$= 4\pi r^2 \left(T - \frac{e^2}{4\pi r^3} \right).$$

Electric charges on drops thus tend to reduce the effective surface tension to $\left(T - \dfrac{e^2}{4\pi r^3} \right)$, so that the vapour pressure of a small charged drop will be smaller than for an uncharged drop. This difference must increase rapidly as the radius of the drop is decreased, so that drops charged with only one electron will have a vapour pressure which is *less* than that over a plane surface when their radius is of the order of 8×10^{-5} cm.

The great importance of this discovery lies in the fact that the path of high speed particles such as α or β radiations can be made visible by the condensation of water vapour on the gaseous ions formed along the track of the rays. For such purposes it is essential that the expansions be performed in such a way that the tracks are completely undisturbed by air currents or eddies. In addition, it is imperative that the expansion be performed and the photograph taken immediately the ions have been produced, and before the tracks can be made nebulous by the normal diffusion occurring in gases.

The Wilson cloud-expansion chamber has been developed into one of the most important instruments which we have available in the study of radioactivity, atomic disintegration, and cosmic rays, and students must refer to original papers and advanced treatises * for accounts of the entrancing discoveries made with the instrument.

§ 23. Low Pressure. Pumps of the Geryk Type

Pumps of this type have a stroke frequency which is very convenient for hand operation. The vacuum obtainable is generally only about $\frac{1}{2}$ mm., but the pump evacuates a container to this pressure with fair speed.

The essential principle of the action of this pump may be gathered from Fig. 24. When the piston descends in the cylinder, the valve closes and a partial vacuum is formed in the space above the piston. Gas from the vessel to be exhausted will flow into this evacuated space when the piston has passed the inlet tube B and will be forced out of the cylinder through the valve on the return of the piston.

The practical pump is more elaborate than this simple model; in order to obviate leakage at piston, piston rod, or valves, and to reduce to zero the 'dead space' round the valve and top of the cylinder which is not traversed by the piston.

FIG. 24

It will be seen from Fig. 25 that the top valve, the piston, and the base of the cylinder are covered with oil, and that the inlet tube B is widened into an annular space to prevent any choking of the tube when oil runs in from the top of the piston.

When the piston rises, its oil layer will expel the air through
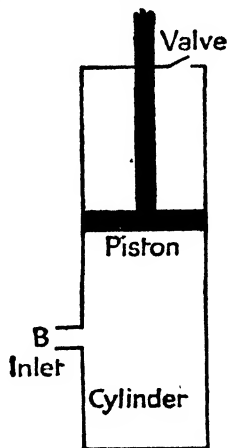
* Such as *Radioactive Substances and their Radiations*, Rutherford, Chadwick, & Ellis.

the valve and will completely fill the cylinder just before the
top of the stroke is reached.   The small quantity of oil which
must be forced through the valve in completing the stroke
will remove from the
vicinity of the valve the
last of the air bubbles.
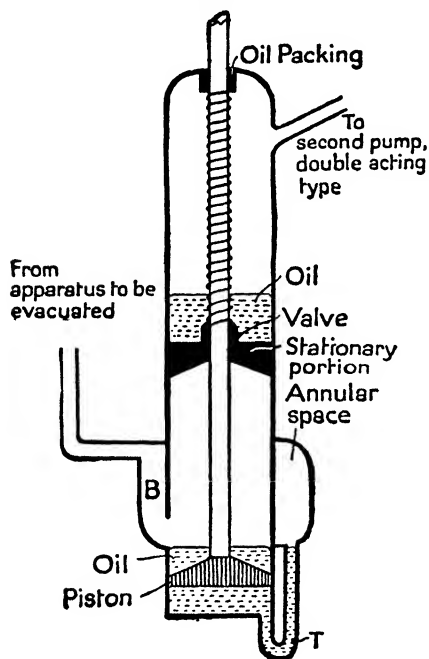
FIG. 25

On the down stroke
the valve will close and
a vacuum will be formed
above the piston.   Any
slight leakage round the
valve or the top of the
piston rod will merely
add to the quantity of
oil above the piston and
will serve to supply the
slight excess desirable
for the next stroke.

When the piston is
nearly at the bottom of
its stroke the hole B
will be uncovered and
the vacuum put into
communication with the
apparatus to be ex-
hausted. Any excess oil
which has run in from
B or has passed the piston will be forced through the
tube T and will be delivered again to the top of the piston.

§ 24. Low Pressure.   Rotary Oil-sealed Pumps

This type of pump was originally developed in connection
with the transference of liquids and is frequently employed
for the oil pump of motor cars, so that useful models demon-
strating their action can be obtained very cheaply from 'car
wreckers.'

Fig. 26 shows (diagrammatically) the interior construction of the first type of these pumps. The outer cylinder A is a heavy casting which is bored and ground with great accuracy. Cylinder B rotates about its central axis so that it is always in contact with the casing at the point marked D. Two plates E and F slide in a slot and are pressed into contact with the inner wall of A by springs G.

On rotating the inner cylinder in the direction of the arrow, the air in space I is compressed and escapes through the valve K. At the same time the volume of space H is increased until the plate F passes the inlet tube J. The air thus entrapped is then carried round the cylinder and is expelled at the valve K.

Fig. 26

The whole of the working parts are placed in an outer rectangular case filled with special oil which ensures perfect lubrication, prevents leakage of air into the high vacuum, and assists in efficient cooling of the pump. A very slight oil leakage into the cylinder along the axles or through the valve is desirable, as it ensures that the surfaces are coated with a thin film of oil which reduces friction and prevents any back-flow of gas at the point D. Surplus oil will be expelled through the valve K.

A type of rotary oil-sealed pump commonly met in laboratories is the 'Hyvac,' which is represented in Fig. 27. In this type the circular rotor is rotated about an axis which
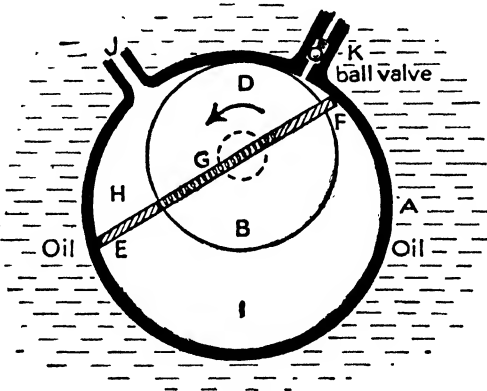
does not pass through its centre, so that, in consequence, the part of the rotor G will rub round the inside of the cylinder, and will always be in contact with it. Between the inlet and the exhaust ports is mounted the movable vane C, which is kept firmly in contact with the rotor by the spring and arm shown at the top of the diagram.
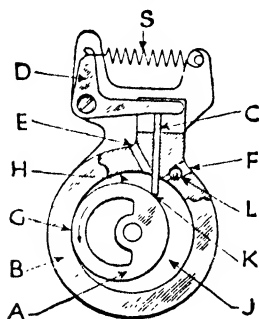
FIG. 27

A. Rotor, eccentrically mounted on shaft passing through centre of:
B. Stator.
C. Vane movable vertically.
D. Vane Pressure Arm.
E. Inlet Port, valveless.
F. Outlet Port.
G. Point of contact between Rotor and Stator.
H. Suction side of Pump.
J. Lead away side of Pump.
K. Point of contact between Rotor and Vane.
L. Steel Ball Exhaust Valve.
S. Tension Spring.

In the complete pumps two of these units are connected in series, and are driven by a common horizontal shaft. The whole is immersed in a metal box filled with special oil of very low vapour pressure. If any leakage occurs along the shaft or vanes of the system, oil alone will be forced into the pump, and will serve—as before —to lubricate the rotor completely. Any excess will be ejected with the air through the ball exit F back into the oil bath.

Rotary oil-sealed pumps of these general types are most conveniently driven by electric motors. They appear to be completely replacing all other types of pumps (except diffusion pumps), since they are relatively small, are silent, robust, and reliable, are easy to operate, and combine a high limiting vacuum with a high speed of pumping.

§ 25. Low Pressure. Diffusion Pumps

Fig. 28 represents a typical single-stage diffusion pump. The mercury in the bulb B is heated by a gas flame or by an

electric heater and boils. The mercury vapour rises up the straight tube T, moves round the bend, past the internal seal S, and condenses in the condenser C; the liquid mercury is then returned to the bulb by the tube D.

The condenser C is cooled by the water jacket W, and the whole of the upright tube T is well lagged with asbestos to prevent the mercury vapour from condensing before it reaches the condenser. The apparatus to be exhausted is connected to the top tube A, while the lower tube B runs to a 'backing' or 'fore' pump which reduces the pressure to the order of 1 mm. of mercury before the diffusion pump can commence operation.

Diffusion pumps differ from the other types so far considered in this essential fact that they can only commence to pump when the pressure has been reduced to between 1 and 5 mm. of mercury. At this pressure the mean free path (p. 49) of hydrogen is of the order of 0·14 mm., and is, therefore, comparable with the dimensions of the apparatus. At this pressure the boiling-point of mercury will be reduced to the order of 150° C. and the volume of 1 gm. of mercury vapour will be so large that any reasonable supply of heat, such as that from a 500-watt resistance unit, will cause the vapour to rush up the tube with the speeds which would be of the order of 3×10^5 cm. a second if the pressure in the boiler were no greater than that of the fore vacuum. The mercury vapour thus moves with speeds which are of the same order as the normal kinetic
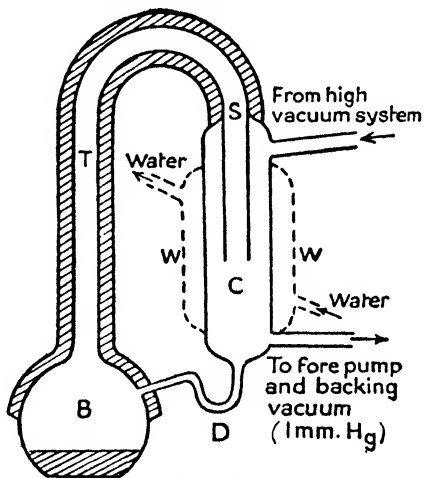
FIG. 28

velocities of ordinary gas atoms ($4\cdot8 \times 10^4$ cm. per sec. at o° C.).

We may obtain a rough physical picture of the operation of the pump by saying that the difference in pressure between the nearly perfect vacuum obtainable in the high vacuum system and the 1 mm. pressure produced by the fore pump is maintained by the destruction of some of the forward momentum of the mercury molecules. Any gas molecules diffusing (i.e. wandering with their ordinary, gas kinetic velocities) back from the fore vacuum will encounter the downward stream of high velocity mercury atoms, and will be swept downwards towards the backing vacuum. Any downwards gas molecules from the high vacuum apparatus that happen to wander into the mercury stream will be knocked by the mercury atoms into the region of the backing vacuum.

Naturally we expect that the mercury atoms will also wander back into the high vacuum system with their ordinary gas kinetic velocities. This does not occur, however, to any great extent, because all that leave the tube and enter the condenser must be moving downwards and can have no upward component of velocity. Moreover, the mercury atoms will very rarely acquire an upward velocity by rebounding from the walls of the tube because the tubes are cold and the vapour pressure of cold mercury is exceedingly small.

Any large concentration of mercury atoms in the condenser immediately above the level of the jet will greatly cut down the speed of pumping because the gas atoms will have to diffuse through the mercury before entering the downward stream of mercury atoms from the jet. It is, therefore, necessary that the level of the water in the cooling jackets should be considerably above the level of the jet.

It is possible with diffusion pumps to produce vacua of 10^{-6} to 10^{-8} mm. of mercury. The real pressure inside the system will not normally be as low as this because some mercury atoms must wander upwards, so that there will be present the vapour pressure of mercury at the temperature of the glass tubes.

This difficulty is overcome by attaching a 'mercury trap' cooled in liquid air (shown in Fig. 29) directly to the tube leading to the high vacuum system.

Diffusion pumps may, therefore, be used to reduce gas pressure to zero from an initial pressure of the order of 1 mm. The high vacuum will be maintained only as long as the mercury stream continues, and if this is stopped for an instant the pressure will immediately become that of the backing vacuum. The pump operates essentially by removing any gas molecules that happen, in their kinetic wanderings, to enter the high speed mercury atom stream. At these low pressures, the mean free path will be so great that molecules will not be impelled from a place of high pressure to one of low pressure by the impact of the other molecules in the high pressure zone as they are at greater pressures.
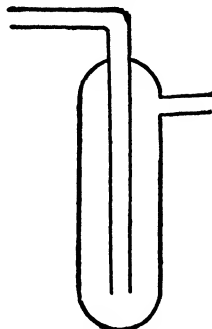
FIG. 29

§ 26. Low Pressure. The Conductance of Tubes and the Speed of Pumps

When the pressure in a tube is so low that the mean free path (i.e. the average distance a molecule moves between collisions) of a molecule is large compared with the dimensions of the apparatus, we may ignore completely the collision of any one gas molecule with any other gas molecule and may assume that molecules move in completely straight lines until they encounter the walls of the containing vessel. The molecules will then not be reflected regularly, but will, instead, attach themselves temporarily to the surface, and will, therefore, be re-emitted in completely random directions, with one half going backwards and one half forwards.

If, therefore, a chamber containing gas at pressures less than 10^{-3} mm. of mercury were to be connected to a perfect

c

vacuum by a long straight tube, the majority of the molecules entering the tube would hit the circumference and would then have an equal probability of going back into the original container at 10^{-3} mm. pressure as of going forward into the perfect vacuum. It is quite clear that the time for the gas to escape into the vacuum would increase very rapidly indeed if the chance of impact with the walls of the tube were increased, by diminishing the diameter D or increasing the length l.

As a result of a very complicated calculation, Knudsen showed that the number of molecules moving through such a tube from a pressure p_1 to a pressure p_2 is proportional to

$$\frac{D^3}{l}\left(p_1 - p_2\right).$$

This result is of vital importance in its practical application, since, as we shall see later, the speed with which a modern diffusion pump will evacuate a container may be altered from days to minutes by replacing 'ordinary sized' connecting tubes by those having large values of D^3.

The speed of a vacuum pump (whose limiting pressure is p_0) is defined by the equation:

$$S = -V\frac{dp}{(p-p_0)dt},$$

where V is the total volume of the container which is being evacuated, p is its pressure, and dp is the decrease in pressure produced by the pump in a time dt.

The speed of modern diffusion pumps is exceedingly great, so that in all ordinary systems the pump can remove gas far more rapidly than the connecting tubes can convey it. Thus mercury diffusion pumps have speeds varying from 2,500 to 150,000* cc. per sec., according to the type and cost, while the speed of a connecting tube 1 metre long and 1 cm. internal diameter is only 100 cc. per sec.

The time necessary to evacuate a 1-litre chamber from

* These speeds are the values quoted by Messrs. W. Edwards & Co. for their diffusion pumps 'K III' and 'Grade K' respectively.

10^{-2} to 10^{-6} mm. of mercury can be calculated with the help of the equation:

$$S = - V\frac{dp}{(p-p_0)dt}.$$

If the two diffusion pumps were connected in turn directly to the chamber, the times required would be 3·7 and 0·06 sec. respectively, whereas if the connecting tube mentioned above were to be employed the time would be increased to 92 secs.

The student may obtain amusement by calculating from the above figures the corresponding times of evacuation when tubes of other lengths and diameters are employed, using the knowledge that the speeds of tubes are proportional to D^3.

§ 27. Low Pressure.   Occluded Gas

The limiting pressure which it is possible to reach in any high vacuum system is profoundly modified by the speed of the vacuum pump and connecting tubes employed.   There are, condensed on the surface and in the pores of vacuum apparatus, quantities of gases and vapours which are liberated only slowly in the highest vacua attainable.   The limiting pressure which can be reached after even a very long time is thus that pressure at which the gas liberated by the surface per second is just equal to the amount which the pump can remove.   It therefore follows that the greater the speed of the pump and its connections the lower will be the limiting pressure reached.

Apparatus in which very low pressures — from 10^{-6} to 10^{-8} mm. of mercury—are to be reached, is made of glass or pyrex, is evacuated, and is then heated in an induction furnace until the glass is near collapsing.   The liberation of occluded gas is not, however, completely prevented even by this drastic treatment, as may be understood when it is recognized that $\frac{1}{10}$ cc. of occluded air would produce 100,000,000,000 cc. at 10^{-8} mm. pressure.

§ 28. Elasticity.   Bodies under Tension

Although discussions of the experimental results of the stretching of wires under direct tension are given in many text-books, students generally have difficulty in understanding the phenomena which occur when the elastic limit is exceeded.
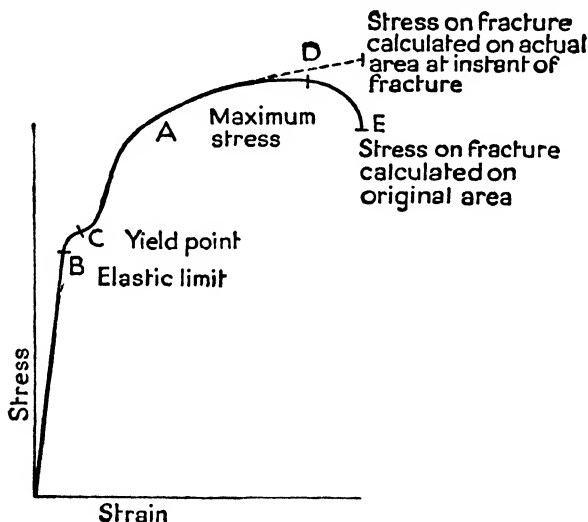
FIG. 30

Fig. 30 represents the type of results which would be obtained in a tensile test with a material with well-defined elastic properties, such as the structural steel used in buildings.

The strain produced in the material—i.e. the ratio $\left(\dfrac{\text{extension}}{\text{original length}}\right)$—is proportional to the stress or tension per unit area until the 'limit of proportionality' is reached.

The constant of proportionality in the relation

$$\frac{\text{force}}{\text{area}} \propto \frac{\text{change in length}}{\text{original length}}$$

is given the name of Young's modulus and the symbol Y (or E),

Thus

$$\frac{\text{force}}{\text{area}} = Y\frac{dl}{l}.$$

At somewhat larger tensions a very considerable increase in the length of the specimen will be produced by a very small increase in the tension per unit area (stress). The point where this occurs is generally called the 'yield point.'. In many materials it is very clearly defined.

Larger forces produce larger extensions until the point D is reached. The ratio of this *maximum force* (i.e. force at D) to the *original area* is called the *ultimate strength* of the material.

Beyond the point D the material will extend rapidly, and, if the force is kept constant, fracture will occur. If, however, the force is reduced, the remainder of the curve beyond D may be followed. This curve beyond D is deceptive because its downward trend is due entirely to the fact that the specimen has now become considerably smaller in cross section so that the ratio

$$\frac{\text{force}}{\text{original area of cross section}}$$

has decreased, whereas the ratio

$$\frac{\text{force}}{\text{actual area of cross section}}$$

has really increased. This is shown by the dotted line.

In this region from D to E a 'waist' generally develops in the material, so that the real force per unit area in the specimen is difficult to calculate and to represent graphically. Moreover, a considerable part of the elongation is produced at this waist, so that it is possible to compare extensions produced in different materials only if the bars are initially of a standard length or if the total elongation can be split up into a general extension proportional to the length and a constant extension characteristic of the waist. This standard length is generally taken as 8 in.

If the force is removed before the limit of proportionality

is exceeded there will be no permanent elongation or *permanent set*. There will, however, certainly be a permanent set if the yield point is exceeded. The exact position on the graph at which the strain (i.e. extension per unit length) ceases to be truly elastic is sometimes called the *true elastic limit* of the material; this limit often—but not always—coincides with the limit of proportionality.

Since the limit of proportionality, the true elastic limit, and the yield point are all near together they are very frequently treated as being one single elastic limit. Most published values and makers' specifications, although called elastic limits, really refer to the yield point, since this is the easiest quantity to measure experimentally.

§ 29. Elasticity. The Time Factor

Time affects stress-strain relationship in two ways. Firstly, the strain, $\left(\dfrac{\text{change in length}}{\text{original length}}\right)$, does not rise instantaneously under the application of the stress, $\left(\dfrac{\text{tension}}{\text{area}}\right)$; it creeps to its final equilibrium value and takes times ranging from fractions of a second to days. Secondly, the position of the yield point and the actual slope of the stress-strain graph (i.e. Young's modulus) varies considerably with the time which has passed since the overstraining occurred.
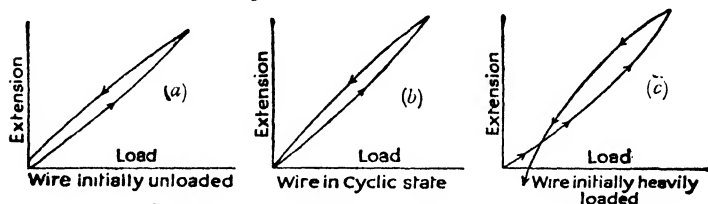
FIG. 31

This slow increase in the strain to its new equilibrium value is very marked with glass fibres, but it is very small with homogeneous substances such as quartz or metals providing

they are not strained beyond the limit of proportionality. Figs. 31 (*a*), 31 (*b*), and 31 (*c*) typify the way in which this creeping of the strain to its equilibrium value modifies experimental stress-strain graphs, even when the elastic limit is not exceeded. The results obtained in normal school or university experiments will gain greatly in regularity if the load is increased uniformly with time. It is also desirable in practical work to bring the wire into a cyclic state before taking the measurements, by loading and unloading it *at the same rate* as the final experimental cycle.

The strength of material is not normally affected by strains which are smaller than the limit of proportionality; the yield point and the Young's modulus of the wire will each be unaffected.

Fig. 32

If the material has been strained to a value between the limit of proportionality and the yield point, an immediate reloading will give a lower limit of proportionality, whereas reloading after a long period of rest will give an increased limit of proportionality. The time necessary for this increased limit to develop is greatly decreased if the material is heated to 100° C. The value of Young's modulus is generally not affected by extensions smaller than the yield point.

If the body is overstrained beyond the yield point and the force is then removed, the decrease in length will correspond to the old value of Young's modulus applying over the whole range. Thus in Fig. 32 BC will be parallel to OA. If the forces are immediately applied again to the body the new Young's modulus is 20% less, so that the body follows the
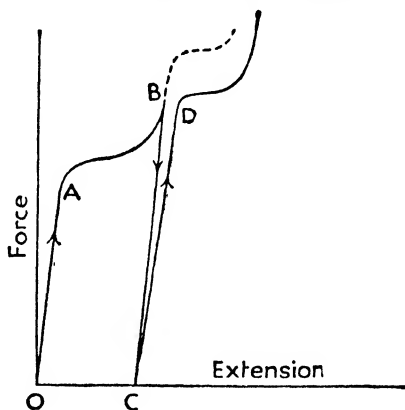
line CD. The original value of Young's modulus will, however, be obtained if the substance is left unloaded for some time, or if it is heated.

Similarly, in this experiment, the new yield point will be the position D if the wire is tested immediately after the removal of the primary load. If the wire is left unloaded for some time, or is heated, the yield point may be much higher than the previous maximum stress. The curve for this is shown by the dotted line.

This increase in the elastic limit of an overstrained wire explains why we can increase the strength of steel wire by some 25% if we add the weights to it very slowly indeed.

The general description of the elastic properties which we have given above applies in particular to structural iron of the type used in engineering construction, bridges, etc. The graphs obtained with other materials are rarely as clear cut; for example, cast iron, stone, and concrete, all of which are weak in tension but strong in compression, have no well-defined elastic limit, and show small amounts of permanent set at low stresses on being first loaded. Glass and other brittle substances, on the other hand, break at stresses which are but very little beyond the elastic limit. In engineering it is highly desirable that the metals employed should be very ductile, because the large extensions which occur after the elastic limit has been passed give adequate warning that the structure is becoming unsafe long before the breaking point is reached. These extensions are, as we have noted, large, and may amount to 20% or even 40% of the original length.

CHAPTER IV

PRISMS, LENSES, AND TELESCOPIC SYSTEMS

§ 30. The Refraction of Light by Prisms

The general method for calculating the angle of deviation produced by a prism on a ray of light is important in itself. This importance is accentuated by the clear insight it affords into the proof and validity of the well-known equation:

$$\mu = \frac{\sin \frac{1}{2}(A + \delta)}{\sin \frac{1}{2}A};$$

and because it gives directly the deviation produced by a prism of small angle, which is the basis of the easiest method of calculating the fundamental equation of lenses.

Consider a ray of light entering a prism at the angle i_1. The ray is deviated at this surface by the amount $(i_1 - r_1)$ and proceeds to strike the second surface at the angle r_2. The total deviation δ produced by the prism is the sum of

that produced at the first surface and that produced at the second surface. Thus:

$$\delta = (i_1 - r_1) + (i_2 - r_2). \qquad (1)$$

In addition to this equation we may note that the angle $(r_1 + r_2)$ is equal to A, since its two arms are normals to the sides of the prism. So

$$A = r_1 + r_2. \qquad (2)$$

In addition to these two equations we know that

$$\sin i_1 = \mu \sin r_1 \qquad (3)$$

and

$$\sin i_2 = \mu \sin r_2. \qquad (4)$$
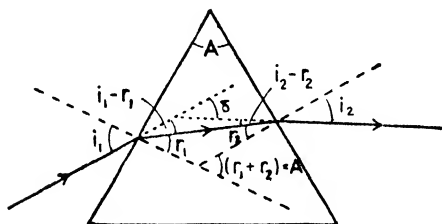
FIG. 33

The easiest way to solve these four equations is to calculate r_1 from equation (3), then r_2 from equation (2), and finally i_2 from equation (4), and then to insert these values in equation (1) to obtain the total deviation.

The procedure is illustrated by the following table, in which each value is obtained from the previous one:

TABLE II

$$\mu = 1\cdot5. \quad A = 60°$$

i_1	$\sin i_1$	$\sin r_1$	r_1	r_2	$\sin r_2$	$\sin i_2$	i_2
30°	0·500	0·333	19° 28	40° 32	0·650	0·975	77·5

$$\delta = (i_1 - r_1) + (i_2 - r_2) = (30° - 19° 28) + (77° 5 - 40° 32)$$
$$= 47° 5.$$

The values of δ shown in the graph (Fig. 34) have all been calculated in the above manner; they range from the maximum possible value of i_1 (90°) to the minimum possible angle at which the ray can emerge from the opposite side. Such a limit must occur because, as i_1 and r_1 become smaller, r_2 increases and must eventually exceed the critical angle, which

means that the ray is totally reflected from this second surface.

Corresponding to each individual value of i_1, such as 70°, for example, there is a particular deviation (45° 51) and a particular value of i_2 (32° 51), and since the path of a ray of light through a prism is reversible, it follows that the same deviation would also occur if t_1 were 32° 51 and i_2 70°. Thus, in
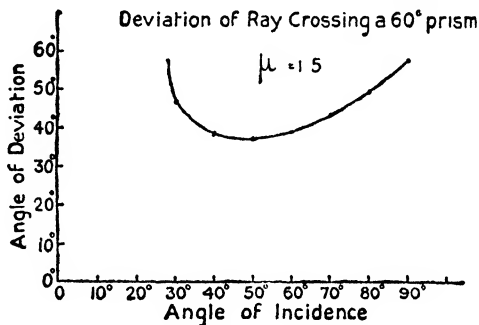
Fig. 34

general, there must exist, as is seen in the graph, two values of i_1 with the same deviation.*

As the deviation approaches the minimum, these two corresponding values of i_1 approach equality, so that at the actual minimum the path of the ray through the prism will be symmetrical and $i_1 = i_2$ and $r_1 = r_2$.

Our four equations for the general refraction of light through the prism then become:

$$\delta = 2(i_1 - r_1) \text{ or } i_1 = \frac{\delta + 2r}{2}$$

$$A = 2r_1$$

and $$\sin i_1 = \mu \sin r_1.$$

* Difficulty is sometimes experienced over the shape of the deviation-angle of incidence graph, and students wonder if curves showing, for example, two maxima and three minima may not occur. The simplest proof that this curve exhibits one minimum only, is a direct experiment carried out of necessity in the process of determining the minimum deviation. Mathematical proof can be produced, however, by equating $\frac{d\delta}{di}$ to 0, proving that $di_1 = -di_2$ at the minimum, and finally expressing the result in terms of r_1 and r_2. Since this proof is quite unimportant and is somewhat tricky, the student is advised not to attempt it.

Therefore $\quad \mu = \dfrac{\sin i_1}{\sin r_1} = \dfrac{\sin \dfrac{A + \delta}{2}}{\sin \dfrac{A}{2}},$

which is the usual formula for prism refraction at minimum deviation.

It is very important to note that this simple formula only applies when the light is traversing the prism at minimum deviation, so that the formula should never be quoted, used, or even considered, without this proviso.

§ 31. Refraction through Prisms of Small Angle

The four general equations for prism refraction give a very simple result for the deviation, if all of the angles involved are so small that the sine of the angle can be taken accurately as the angle in radians. The two interior angles, r_1 and r_2, can, however, only be small simultaneously if the angle of the prism also is small. The simplification is, therefore, only possible if the rays are traversing small angle prisms at angles which are nearly normal to the surfaces. The four equations then become in their simplified form:

$$\delta = (i_1 - r_1) + (i_2 - r_2) \qquad A = r_1 + r_2$$
$$i_1 = \mu r_1 \qquad\qquad i_2 = \mu r_2$$

so that we have:

$$\delta = i_1 + i_2 - A = \mu A - A$$
and $$\delta = (\mu - 1) A*$$

This simple result is of particular importance from the fact that the deviation is independent of the angle of incidence over a wide range. The extent of the validity of this approximate formula can be gathered from the following table,

* Text-books sometimes obtain this formula by direct approximate from the formula for minimum deviation. This is clearly quite wrong since the formula for small angle prisms is valid for a finite range of i_1, and not merely for the particular angle for which the rays are symmetrical.

which has been calculated accurately. It should be observed in this connection that the smaller the angle of the prism the more independent will the deviation be of the angle of incidence.

TABLE III

$\mu = 1\cdot5. \quad A = 10°$

i_1	$-30°$	$-20°$	$-10°$	$0°$	$10°$	$20°$	$30°$	$40°$	$50°$
δ	$7°\,47$	$6°\,11$	$5°\,27$	$5°\,6$	$5°\,2$	$5°14$	$5°32$	$6°32$	$7°57$

§ 32. The Theory of Thin Lenses

The simple statement that the deviation produced by a small angle prism is constant and is approximately independent of the angle of incidence of the ray, can be used to deduce the normal object and image relation of a lens which has such a small aperture that the tangents of the angles are accurately equal to the angles themselves.
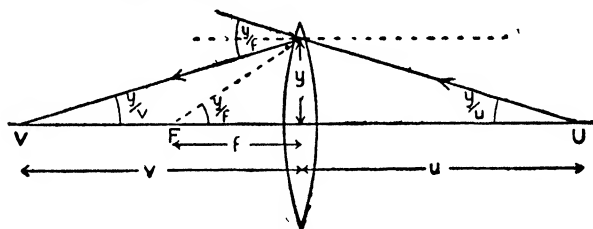
FIG. 35

Regard the lens as a collection of prisms of varying angle which are arranged together to deviate parallel rays incident on them so that they meet together at a point called the principal focus. Such a result can be obtained if the angle of the prism, and therefore the deviation, increases as the prism gets further from the optic axis.

Now the deviation, $\dfrac{y}{f}$, imposed on parallel rays is independent of the angle of incidence, and so will be the same for a ray coming from an object at a distance u and going

to an image at a distance v. Fig. 35 shows that this deviation is, however, $\dfrac{y}{v} + \dfrac{y}{u}$; so we have:

$$\frac{y}{f} = \frac{y}{v} + \frac{y}{u},$$

and we obtain the usual formula for a lens:

$$\frac{1}{f} = \frac{1}{v} + \frac{1}{u}.$$

§ 33. The Conventions of Signs concerning Lenses and Mirrors

The simple formula we have just obtained is correct if it is used in conjunction with its correct sign convention. There are, however, eight possible conventions of signs, of which at least three are in common use, and it is best to define these before attempting to compare their merits and utilities.

(a) $\dfrac{1}{f} = \dfrac{1}{v} + \dfrac{1}{u}$. In the convention of signs for which this equation is correct; a real object always has a positive value of u, while a positive value of v means that the image is real and is on the side of the lens or the mirror in which real images are found (i.e. with a lens, the image is on the side opposite to the real object). Convex, i.e. converging, lenses are counted as having a positive focal length, since they have a real focus; while concave, i.e. diverging, lenses have a negative focal length.

(b) $\dfrac{1}{f} = \dfrac{1}{v} - \dfrac{1}{u}$. In this convention of signs the lengths v and u are measured from the lens, but are counted positive if, in order to reach the object or image considered, the observer has to move *in the same direction* as that of the incident light. A simple development of this gives the final convention that convex, i.e. converging, lenses are counted as having a positive focal length, while concave, i.e. diverging, lenses have a negative focal length.

(c) $\dfrac{1}{f} = \dfrac{1}{v} - \dfrac{1}{u}$. In the *second* convention of signs for which this equation is correct the lengths v and u are measured from the lens, and are counted positive if, in order to reach the image or object, the observer moves in a *direction opposite* to that of the incident light. A simple development of this gives the further convention that convex, i.e. converging, lenses are counted as having a negative focal length, while concave, i.e. diverging, lenses have a positive focal length.

§ 34. Comparison of the Conventions of Signs

The most important single item regarding the conventions of signs is that a student must *thoroughly* understand one convention and be completely free from confusion. Experience shows that very few people indeed have a clear enough grip of their conventions of signs to withstand the excitement of examination, the haze of argument, or the passage of years. The student is, therefore, well advised to cross-examine himself on every opportunity until the correct assignation of signs in such calculations becomes a matter more of intuition than of painstaking thought.

Most students reading this book in the immediate future will have met the last-named convention of signs, as this is, at present, given in most school text-books. It seems, however, probable that either convention (*a*) or convention (*b*) will in the course of time completely replace the last, and there are, at the moment, definite attempts being made to reach some uniformity regarding sign conventions in general and, in particular, regarding the sign denoting the difference between a convex and a concave lens. The whole of the optical industry is irrevocably fixed to the convention of attaching a positive sign to a convex lens; it is quite futile to hope that industry will change, and so if ever one orders a lens of $+15$ cm. focal length one will inevitably obtain a convex lens. If, therefore, any unification is to be achieved at all, schools and students must forget their conventions and

must come into line with the rest of the world and call a convex lens positive. The author, therefore, entreats each progressive student of physics personally to help in making the subject clear and intelligible, to forget his former convention, and to learn one of the two normal conventions.

Of the two conventions (*a*) and (*b*) in which convex lenses are counted positive, the author definitely prefers the first $\left(\dfrac{1}{f} = \dfrac{1}{v} + \dfrac{1}{u} \right)$, in view of the fact that most students will have to unlearn one convention and a more radical change seems likely to lead to less confusion than would the subtler change from measuring in the same or in opposite directions to the incident light. In addition, his own experience (and that of many other teachers who have tried both systems) shows that students make fewer mistakes with this convention than with the others. This is, however, an entirely personal preference, since the Committee of the Physical Society, which has recently reported on the matter, was unable to decide unanimously which system was the better.

Throughout the following sections this sign convention $\left(\dfrac{1}{f} = \dfrac{1}{v} + \dfrac{1}{u} \right)$ **is uniformly adopted, but whenever there is difficulty, the treatment is also carried out with the other conventions.**

In view of the existence of all these different conventions of signs, each student is advised to state at the beginning of every examination answer to a question on lenses, the precise sign convention he is employing.

The convention of signs for which the formula $\dfrac{1}{f} = \dfrac{1}{v} + \dfrac{1}{u}$ is applicable to both lenses and mirrors may now be repeated.

Lenses or mirrors which make parallel rays converge to a focus are said to have a positive focal length. Lenses or mirrors which make parallel rays diverge from a virtual focus are said to have a negative focal length.

The surface of a lens which separates media of smaller refractive index (i.e. a glass lens in air) is said to have a positive radius of curvature if it makes incident parallel rays converge, and is said to have a negative radius of curvature if it makes incident parallel rays diverge.

The signs of these radii remain the same if the lens is immersed in media of greater refractive index, and the negative character of the focal length of a diverging biconvex air lens surrounded with water of greater refractive index is shown by the negative value of the term

$$\left(\frac{\mu_{lens}}{\mu_{medium}} - 1 \right)$$

Distances with either lenses or mirrors which are actually traversed by the light are counted positive, while those only virtually covered are negative.

For Lenses—irrespective of the Direction of the Light—

$$\frac{1}{v} + \frac{1}{u} = \frac{1}{f} = \left(\mu - 1 \right) \left(\frac{1}{r_1} + \frac{1}{r_2} \right)$$

Real objects have positive values for u.
Real images have positive values for v.
Ordinary convex lenses in air have positive focal lengths.
The radii of curvature of ordinary biconvex lenses are positive.

Virtual objects have negative values for u.
Virtual images have negative values for v.
Ordinary concave lenses in air have negative focal lengths.
The radii of curvature of ordinary biconcave lenses are negative.

Meniscus lenses have one radius positive and one negative.

For Mirrors—irrespective of the Direction of the Light—

$$\frac{1}{v} + \frac{1}{u} = \frac{1}{f}$$

Real objects have positive values for u.

Real images have positive values for v.

Ordinary concave mirrors have positive radii of curvature and positive focal lengths.

Convex mirrors have negative radii of curvature and negative focal lengths.

§ 35. Calculation of the Focal Length of Thin Lenses

On page 61 we have considered a thin lens as being made up of an infinite collection of small angle prisms whose angle increases with the distance from the centre of the lens in such a way that parallel rays are all brought together at a point called the principal focus. We will now show that this will occur if the surfaces are portions of spheres.
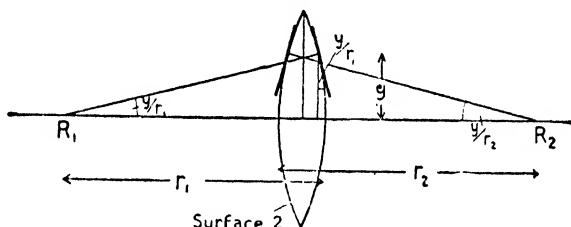
FIG. 36

The angle between the tangent and the vertical is $\dfrac{y}{r_1}$ in radian measure, if the angle is small enough for tan θ to be accurately equal to θ (Fig. 36).

The effective angle of the prism at the height of y is then

$$\left(\frac{y}{r_1} + \frac{y}{r_2}\right).$$

The deviation imposed on a ray of light is therefore

$$\left(\mu - 1\right)\left(\frac{y}{r_1} + \frac{y}{r_2}\right),$$

since this deviation is $(\mu - 1)$ (the angle of the prism) and this

must equal the deviation given to a ray of light incident parallel to the axis, so

$$\frac{y}{f} = \left(\mu - 1\right)\left(\frac{y}{r_1} + \frac{y}{r_2}\right).$$

In this equation the quantity y cancels completely, so that the distance from the lens (f) at which parallel rays cross the axis after refraction is independent of y, and must, therefore, be the same for all rays whatever may be the distance from the centre. Thus such a lens with spherical glass surfaces produces a point image from parallel rays and

$$\frac{1}{f} = \left(\mu - 1\right)\left(\frac{1}{r_1} + \frac{1}{r_2}\right). *$$

In the convention of signs $\frac{1}{v} - \frac{1}{u} = \frac{1}{f}$, positive in same direction as the incident light; if the light is considered to be moving in the direction $\longleftarrow$ r_1 will be in the direction of the light and so will be positive, while r_2 will be negative, so that

$$\frac{1}{f} = \left(\mu - 1\right)\left(\frac{1}{r_1} - \frac{1}{r_2}\right).$$

In the convention of signs $\frac{1}{v} - \frac{1}{u} = \frac{1}{f}$, positive opposite to the direction of incident light; if the light is considered moving in the direction $\longleftarrow$ r_1 will be in the direction of the light and so will be negative, so that

$$\frac{1}{f} = \left(\mu - 1\right)\left(\frac{1}{r_1} - \frac{1}{r^2}\right),$$

since f is also negative.

* The equation obtained above applies only for a material lens in a medium of refractive index unity. If the lens is immersed in a medium of refractive index μ_m the formula becomes:

$$\frac{1}{f} = \left(\frac{\mu}{\mu_m} - 1\right)\left(\frac{1}{r_1} + \frac{1}{r_2}\right).$$

This more general and symmetrical equation follows directly from the generalized deviation equation for a small angle prism:

$$\delta = \left(\frac{\mu}{\mu_m} - 1\right)\left(A\right).$$

This method for obtaining the above formula should be regarded as being additional to the normal ray method and, in particular, to the wave method, which are adequately discussed in the usual text-books.

§ 36. Combination of Thin Lenses, out of Contact

Consider a ray incident parallel to the axis (Fig. 37); it will be deviated by the first lens through an angle $\frac{y}{f}$, it will then proceed and strike the second lens (f_2) at a height $\left(y - \frac{ay}{f_1}\right)$ above the axis. This second lens, like any other lens, produces
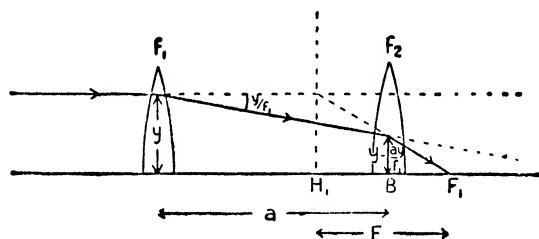
FIG. 37

on any ray, whatever its angle, a deviation equal to $\frac{\text{height above axis}}{f_2}$. The second lens, therefore, produces an additional deviation of $\dfrac{\left(y - \dfrac{ay}{f_1}\right)}{f_2}$, so that the final emergent ray, therefore, makes an angle

$$\frac{y}{f_1} + \left(\frac{y}{f_2} - \frac{ay}{f_1 f_2}\right)$$

with the axis.

If we had wished to use one thin lens alone to make the rays of light traverse the same path that they do *after they have emerged* from the second lens, we should have had to

place the single lens at H_1 and to employ a lens whose focal length was H_1F_1.* This length H_1F_1 is called the equivalent focal length of the lens combination (F), and the point H_1 is called a principal point of the combination.

The angle $\left(\dfrac{y}{f_1} + \dfrac{y}{f_2} - \dfrac{ay}{f_1f_2}\right)$ is clearly also $\dfrac{y}{F}$, so

$$\frac{1}{F} = \frac{1}{f_1} + \frac{1}{f_2} - \frac{a}{f_1f_2}.$$

This equation, therefore, permits us to calculate the equivalent focal length F of the combination of two thin lenses separated by a distance a.

If required, we can easily calculate the length H_1B—the distance from the principal plane to the second lens. It is simply the length $\dfrac{ay}{f_1}$ divided by the angle $\dfrac{y}{f_1} + \dfrac{y}{f_2} - \dfrac{ay}{f_1f_2}$, which is $\dfrac{a}{1 + \dfrac{f_2}{f_1} - \dfrac{a}{f_2}}$.

In the convention of signs $\dfrac{1}{v} - \dfrac{1}{u} = \dfrac{1}{f}$, in which the positive is in the same direction as the incident light, both f_1 and f_2 are positive since they are converging lenses; the angles of deviation $\dfrac{y}{f_1}$, etc., are all positive angles, and the length a is measured from the first to the second lens and is again positive. Thus the length H_1F_1, which is, of course, F, the equivalent focal length of the combination, is given by the equation:

$$\frac{1}{F} = \frac{1}{f_1} + \frac{1}{f_2} - \frac{a}{f_1f_2}.$$

In the convention of signs $\dfrac{1}{v} - \dfrac{1}{u} = \dfrac{1}{f}$, in which the positive is in the opposite direction to the incident light, both f_1 and f_2 are negative since they are converging lenses, the angles of deviation

* Students must refer to the theory of thick lenses before attempting to employ this theory for rays which are other than parallel to the axis.

$\frac{y}{f_1}$, etc., are all negative angles, the length a may be negative or positive according to whether it is measured from the first to the second lens or from the second to the first lens.   Thus the length H_1F_1, which is, of course, $-F$, is given by:

$$-\frac{1}{F} = -\frac{1}{f_1} - \frac{1}{f_2} \pm \frac{a}{f_1f_2}.$$

§ 37. Lenses in Contact

The effective focal length of two lenses in contact is obtained directly by inserting the value $a = 0$ in the equation for lenses out of contact, and so

$$\frac{1}{F} = \frac{1}{f_1} + \frac{1}{f_2}$$

suffices for all conventions.
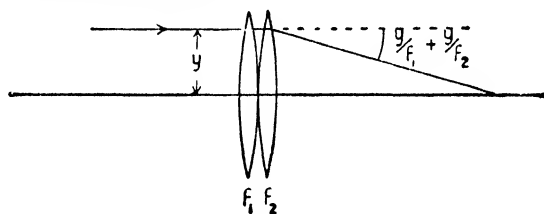
FIG. 38

If a separate proof is required, we may note that the deviation (shown in Fig. 38) produced by the first lens is $\frac{y}{f_1}$ and that produced by the second lens is $\frac{y}{f_2}$, since the height of the ray above the axis is also y for the second lens if they are in contact.   The total deviation is thus $\frac{y}{f_1} + \frac{y}{f_2}$, and this is also $\frac{y}{F}$.   Thus

$$\frac{1}{F} = \frac{1}{f_1} + \frac{1}{f_2}.$$

The ratio $\frac{1}{f}$ for a lens or a mirror is frequently spoken of as the 'power' of the lens. This is in accordance with the popular use of the term in which a powerful lens is one with a short focal length. The above equation can therefore be expressed by the statement that the resultant power of a number of thin lenses in contact is the algebraic sum of the powers of the separate components.

§ 38. Achromatic Lenses

Since the focal length of lenses varies with refractive index according to the equation:

$$\frac{1}{f} = (\mu - 1)\left(\frac{1}{r_1} + \frac{1}{r_2}\right),$$

and since for all substances (μ_{blue}) is greater than (μ_{red}), any single lens will have a greater focal length for the red than for the blue. In consequence, images formed by such a lens will not be sharp, but will exhibit coloured margins.

It is possible, however, to construct achromatic compound lenses, consisting of two, three, or more single lenses of different glass placed in contact in such a way that the focal lengths for two, three, or more colours respectively are the same. The most important practical example of this uses two lenses, of flint and crown glass respectively.

In ordinary practical problems, one knows the refractive indices μ_r', μ'_b, μ''_r, μ''_b for red and blue rays of each of the types of glass from which it is proposed to construct the achromatic combination, and, in addition, one knows the focal length F of the combination it is desired to construct. The unknowns are the radii of curvature of the four surfaces of the two lenses, r_1', r_2', r_1'', and r_2''. To evaluate these four unknowns we need ultimately four equations. The first of these expresses the fact that the focal length

of the combination is the same for red and for blue rays, i.e.:

$$\frac{1}{F_b} = \frac{1}{F_r},$$

or

$$\frac{1}{f'_b} + \frac{1}{f''_b} = \frac{1}{f'_r} + \frac{1}{f''_r}.$$

And therefore:

(1)
$$\left(\mu'_b - 1\right)\left(\frac{1}{r_1'} + \frac{1}{r_2'}\right) + \left(\mu''_b - 1\right)\left(\frac{1}{r_1''} + \frac{1}{r_2''}\right)$$

$$= \left(\mu'_r - 1\right)\left(\frac{1}{r_1'} + \frac{1}{r_2'}\right) + \left(\mu''_r - 1\right)\left(\frac{1}{r_1''} + \frac{1}{r_2''}\right).$$

The second equation expresses the fact that the focal length of the combination must be the desired value F, and so:

(2)
$$\left(\mu'_b - 1\right)\left(\frac{1}{r_1'} + \frac{1}{r_2'}\right) + \left(\mu'_b - 1\right)\left(\frac{1}{r_1''} + \frac{1}{r_2''}\right) = \frac{1}{F}.$$

(3) The third equation is merely that $r_2' = -r_1''$, since it is usual to unite the two lenses with a film of Canada balsam, partly for convenience, and partly in order to diminish loss of light by reflection at the surfaces.

(4) The fourth equation expresses the fact that one of the surfaces has some particular curvature that is convenient, such as, for example, $r_2'' = \infty$. The curvature of this surface is frequently chosen in order to reduce the spherical aberration to a minimum for the particular objects and images for which the lens is designed.

By simple algebraic manipulation of the above equations all problems on twin achromatic lenses can be solved if, in addition, it is remembered that the quantity $\dfrac{\mu'_b - \mu'_r}{\mu'_{mean} - 1}$ is called the dispersive power of glass (v), so that if one is given μ_{mean} and the dispersive power v, one must first calculate μ_b and μ_r.

The arithmetical part of the calculation is frequently made

easier at the expense of the clarity of the expression of the physical principles by reducing* equation (1) to

$$\mu'\text{b}\left(\frac{1}{r_1'}+\frac{1}{r_2'}\right) + \mu''\text{b}\left(\frac{1}{r_1''}+\frac{1}{r_2''}\right) = \mu'\text{r}\left(\frac{1}{r_1'}+\frac{1}{r_2'}\right) + \mu''\text{b}\left(\frac{1}{r_1''}+\frac{1}{r_2''}\right),$$

and then to

$$\left(\mu'\text{b}-\mu'\text{r}\right)\left(\frac{1}{r_1'}+\frac{1}{r_2'}\right) + \left(\mu''\text{b}-\mu''\text{r}\right)\left(\frac{1}{r_1''}+\frac{1}{r_2''}\right) = 0,$$

and so to

$$\left(\mu'\text{b}-\mu'\text{r}\right)\left(\frac{\mu'm-1}{\mu'm-1}\right)\left(\frac{1}{r_1'}+\frac{1}{r_2'}\right) +$$

$$\left(\mu''\text{b}-\mu''\text{r}\right)\left(\frac{\mu''m-1}{\mu''m-1}\right)\left(\frac{1}{r_1''}+\frac{1}{r_2''}\right) = 0$$

and to

$$\frac{v'}{f'} + \frac{v''}{f''} = 0.$$

It is interesting to note that this shows that achromatic doublets can only be constructed from combinations of converging and diverging lenses since dispersive powers may vary in magnitude but never vary in sign.

The treatment given in this section can instantly be adopted to refer to the other two conventions by replacing the term $\left(\frac{1}{r_1} + \frac{1}{r_2}\right)$ everywhere by $\left(\frac{1}{r_1} - \frac{1}{r_2}\right)$.

* This is simplified if we employ the methods of the calculus.

$$\frac{1}{f'}+\frac{1}{f''} = \frac{1}{F} = \text{a constant.}$$

$$\therefore \delta\left(\frac{1}{f'}\right) + \delta\left(\frac{1}{f''}\right) = 0.$$

But $$\delta\left(\frac{1}{f'}\right) = \delta\mu'\left(\frac{1}{r_1'} + \frac{1}{r_2'}\right) = \frac{\delta\mu'}{\mu'-1}\left(\frac{1}{f'}\right) = \frac{v'}{f'}.$$

So the condition for achromatism is that $\frac{v'}{f'} + \frac{v''}{f''} = 0.$

§ 39. Internal Reflections in Lenses

Consider a ray (Fig. 39) coming from an object and then being reflected from the far surface of the lens, which will reflect a little light even though it is not silvered. The rays finally emerge from the lens after reflection and form an image
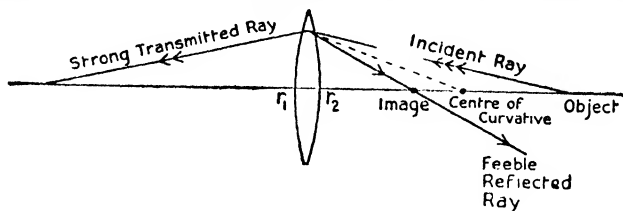
FIG. 39

just as they would have done if they had been reflected from a single concave mirror of an equivalent focal length F. We can easily find this equivalent focal length by adding the powers of the lens and the powers of the mirrors, regarding them all as being in contact. (On page 70 we proved that $\frac{1}{F} = \frac{1}{f_1} + \frac{1}{f_2}$ if the lenses are in contact, and this holds also for mirrors.) Thus, as we trace the ray coming from the object, we must include first the power of the lens $\left(\frac{1}{f}\right)$ as the ray traverses it before reaching the reflecting glass surface, then the power of this reflecting glass mirror $\left(\frac{2}{r_1}\right)$ as the light is being reflected, and finally the power of the lens again as the ray traverses it before emerging. Thus:

$$\frac{1}{F} = \frac{1}{f} + \frac{2}{r_1} + \frac{1}{f}.$$

We can always understand why the power of the lens should be included a second time, by considering the reflecting surface separated from the second surface of the lens by an infinitely thin air film. Further confirmation is obtained by noting that the reason why a lens brings parallel

rays to a focus is because its middle is thicker than its circular edge, so that the rays traversing the middle are delayed more than those traversing the periphery. In our example both the middle and the periphery are traversed twice, so that the deviation produced by the lens is increased and the focal length doubled.

The general principle of the above treatment can easily be extended. Trace a ray from the object, and add in the power of every lens it traverses and of every surface it is reflected from, counting converging lenses or surfaces positive. It can be illustrated by the following example (Fig. 40).
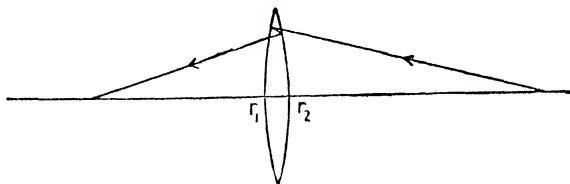
FIG. 40

$$\frac{1}{F} = \frac{1}{f} + \frac{2}{r_1} + \frac{1}{f} + \frac{2}{r_2} + \frac{1}{f} = 2\left(\frac{1}{r_1} + \frac{1}{r_2}\right) + \frac{3}{f}.$$

$$\therefore \frac{1}{F} = \left(\frac{2}{\mu - 1}\right)\frac{1}{f} + \frac{3}{f},$$

$$\therefore F = f\frac{\mu - 1}{3\mu - 1},$$

and if we make the rough approximation that $\mu = 1\cdot5$,

$$F = f\frac{0\cdot5}{3\cdot5} = \frac{f}{7}.$$

This particular type of problem is not so conveniently dealt with by either of the other two conventions, since the rays of light are continually altering in direction.

§ 40. The Theory of Thick Lenses

The term 'thick lens' is a very general one, and is used to refer not only to single lenses, but also to any combination of

lenses which happen to be mounted in such a way that the individual components cannot be used separately. Such combinations are nearly always mounted so that the optic centres of the components are all arranged on one straight line.

We know that each of the surfaces and each of the lenses constituting the thick, compound lens will form an image of an object, hence rays from an object on one side of the lens will emerge on the other, converging to—or diverging from—an image.

Now consider the incident ray which is coming from the left parallel to the axis; when it strikes the first surface of
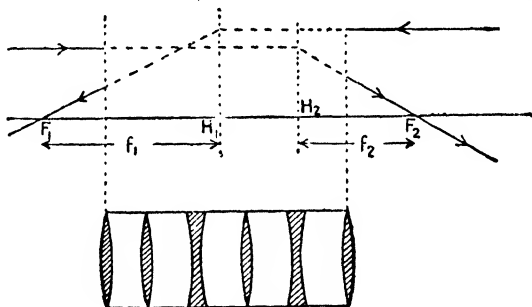
FIG. 41

the thick lens it becomes inaccessible to direct observation, but on emerging from the other side its path is as shown in Fig. 41.

The point where it passes through the axis F_2 must be the focus of the lens to which all incident rays parallel to the axis must be converging, since we know the lens must really have a focus and since two rays meet at that point. The point H_2 is clearly important, since, if we had been trying to make the emerging ray travel exactly the same path after deviation by a thin lens, we should have had to place the lens at H_2, and we should have had to choose the distance H_2F_2 as its focal length.

The plane through H_2 is called the principal plane of the thick lens and f_2 is the focal length.

In a similar way, rays incident from the right form the points F_1 and H_1, separated by f_1.

Notice that the points H_1 and H_2 may be as shown, may be inverted, or may even be exterior to the lens. If the rays were diverging from the lens after refraction, the point F_2 would have been inside the lens, and the sign of the lens (converging or diverging) would have depended on which side of F_2 was the point H_2.

We have now found that rays which are coming from the left and which are parallel to the axis emerge as if they had been bent at the plane through H_2, so that they pass through F_2.

Now consider two rays coming from an object (Fig. 42); the
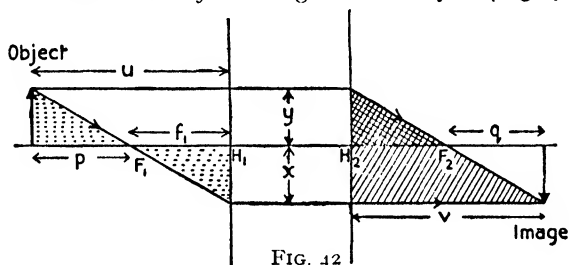
FIG. 42

particular one which was initially parallel to the axis emerges as shown, while the ray which initially passed through the focus F_1 emerges on the other side as if it had been made parallel to the axis at the plane through H_1 (rays must be reversible). *Notice that the rays are not really bent at H_1 or H_2. They are actually bent at the surfaces of the lenses making up the thick lens, but, after emerging from the glass, they behave as if they had been bent at the two principal planes.*

The position where the two rays intersect must be the position of the image, since all rays intersect at the image, and since two straight lines can only intersect at one point.

Now consider the large single-shaded triangle and the smaller double-shaded one; clearly

$$\frac{y}{x+y} = \frac{f_2}{v}.$$

Similarly, from the two corresponding triangles on the other side (*not* the dotted ones),

$$\frac{x}{x+y} = \frac{f_1}{u},$$

so that, adding these two equations,

$$\frac{f_1}{u} + \frac{f_2}{v} = 1.$$

This result is precisely the same as that obtained for thin lenses, because we can prove that $f_1 = f_2$, unless one side of the lens is immersed in water or some medium other than air. The only difference is that in the equation $\frac{1}{u} + \frac{1}{v} = \frac{1}{f}$, the distance u is measured from the principal plane H_1 and the distance v from the principal plane H_2, and these planes will not coincide with the glass surfaces of the lens and may even be completely exterior to them.

In the convention of signs in which the positive is in the same direction as the incident light, and in which f is positive for a convex lens, both f_1 and f_2 will be positive, the length H_2I will be negative and so will be called $+v$, but the length H_1O will be positive and will be called $-u$, so we have the formula:

$$\frac{1}{v} - \frac{1}{u} = \frac{1}{f}.$$

In the convention of signs in which the positive is in the opposite direction to the incident light and in which f is negative for a convex lens, both H_1F_1 and H_2F_2 will be negative and will be called $-f_1$ and $-f_2$ respectively.

The length H_2I will be negative and will be called $-v$, and the length H_2O will be positive and will be called $+u$, so we obtain the formula:

$$\frac{1}{v} - \frac{1}{u} = \frac{1}{f}.$$

A further interesting equation may be proved by considering

the two dotted triangles, in which $\frac{x}{y} = \frac{f_1}{q}$, and the two corresponding triangles on the other side of the diagram $\frac{x}{y} = \frac{q}{f_2}$.
Thus, combining the two:

$$pq = f_1 f_2.$$

This is Newton's formula for a thick lens, and is precisely the same as the little-known Newton's formula for a thin lens, where the distances p and q are measured from two foci respectively.

§ 41. Experimental Work with Thick Lenses

Practical work with 'thick lenses' usually involves either the determination of the positions of the points F_1, F_2, H_1, and H_2, (Fig. 43) with respect to the surface of the lens, or the testing of the theoretical formula derived above.
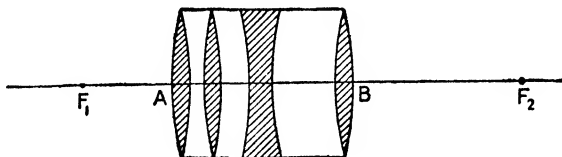
FIG. 43

The student is advised to commence his experiment by drawing an optic axis and marking on it, *to scale*, the position of the lens surfaces. Next it is convenient to place a plane mirror on one side and to adjust a pin on the other till it is in parallax with its own image. The point F_1 is thus obtained. The procedure is then repeated on the other side and the point F_2 obtained.

The focal length of the thick lens is now found by the application of Newton's formula, $pq = f^2$, using a pin on one side at a measured distance p from that focus, and adjusting another pin on the other side to be in parallax with the image and noting its distance q from the focus on that

side. On the scale diagram the distance f is marked off from F_1 and F_2 respectively, and the positions of H_1 and H_2 inserted.

§ 42. Calculation of Principal Planes for a Single Thick Lens of Small Aperture
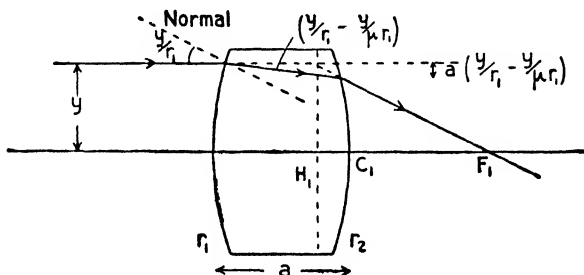
FIG. 44

In this calculation all of the angles involved are so small that the sines of the angles are equal to the angles themselves in radian measure. We further assume that the aperture of the lens is so small that the thickness (a) is sensibly constant throughout, and that the ray strikes the second surface at approximately the same height (y) as it struck the first surface. These assumptions are not so foolish as they seem, because if they were not approximately true a simple lens would probably not form a sharp image, and we are interested only in lenses which do give a sharp image.

The incident ray makes an angle $\dfrac{y}{r_1}$, with the normal to the first surface. After refraction its angle with the normal is $\dfrac{y}{\mu r_1}$, so it has been deviated by $\left(\dfrac{y}{r_1} - \dfrac{y}{\mu r_1}\right)$. Its angle with the normal to the second surface will then be approximately

$$\frac{y}{r_2} + \left(\frac{y}{r_1} - \frac{y}{\mu r_1}\right).$$

After refraction its angle with this normal will be

$$\mu\left\{\frac{y}{r_2} + \left(\frac{y}{r_1} - \frac{y}{\mu r_1}\right)\right\}$$

so that the angle of the ray with the horizontal is:

$$\mu\left\{\frac{y}{r_2} + \left(\frac{y}{r_1} - \frac{y}{\mu r_1}\right)\right\} - \frac{y}{r_2}.$$

In the theory of thick lenses discussed in the previous section the formula $\frac{1}{u} + \frac{1}{v} = \frac{1}{f}$ applied accurately providing the distances u and v were measured respectively from the principal points H_2 and H_1. This formula will, therefore, apply accurately for the single thick lens we are now considering, if the distances are also measured from the principal points H_1 and H_2.

We therefore wish to calculate the distance H_1C, so that if we are using the formula $\frac{1}{u} + \frac{1}{v} = \frac{1}{f}$ we can obtain the true value of u and v by adding the length H_1C_1 to the observed distance from the lens front. This length H_1C_1 is clearly the vertical distance $a\left(\frac{y}{r_1} - \frac{y}{\mu r_1}\right)$ divided by the angle the emergent ray makes with the horizontal (see Fig. 44). Thus:

$$H_1C_1 = -\frac{a\left(\dfrac{y}{r_1} - \dfrac{y}{\mu r_1}\right)}{\mu\left\{\dfrac{y}{r_2} + \left(\dfrac{y}{r_1} - \dfrac{y}{\mu r_1}\right)\right\} - \dfrac{y}{r_2}}$$

$$= \frac{a}{\mu + \dfrac{\mu\dfrac{y}{r_2} - \dfrac{y}{r_2}}{\dfrac{y}{r_1} - \dfrac{y}{\mu r_1}}}$$

$$= \frac{a}{\mu + \mu\dfrac{r_1}{r_2}}$$

D

Now most lenses normally encountered in laboratory work are biconvex and have $\mu = 1.5$. For such lenses the correction H_1C_1 to be added to every experimental value of the distance from object or image to surface of lens in order to make it u or v is $\dfrac{a}{3}$.

A further type of lens frequently encountered is the planoconvex, and for this the correction is zero for distances measured from the plane side, and $\dfrac{a}{1.5}$ for distances measured from the convex side.

§ 43. Defects of Images

Lenses are, in general, used to produce images, and normally exhibit a number of 'defects' when used for this purpose. These defects are:

(1) *Spherical aberration*, in which rays of one single wave length proceeding from a point object do not exactly meet at a point. Lenses can now be made by compounding several separate lenses of special kinds of glass with carefully designed curvatures so that the spherical aberration is very small or is zero *when the object and image are at the particular distances for which the lens was designed.*

(2) *Distortion.* Even if a lens has no spherical aberration, so that a point object produces a point image, a distortion of a line object can occur; for the magnification produced by a lens is dependent on the distance of the different parts of the line object from the optic axis. The distortion produced by a given lens can be materially altered by introducing a diaphragm or stop in front, in the middle, or at the back of the lens.

(3) *Chromatic aberration.* This has already been discussed in the paragraph on achromatic lenses.

(4) *Curvature of field.* It is an experimental fact that the images produced by a simple lens, of points off the axis of

the lens, do not all lie in a plane perpendicular to the axis. This is illustrated in Fig. 45, from which it will be noted that *two* images are produced when the object is well off the axis. These images are not points, but are lines, the one
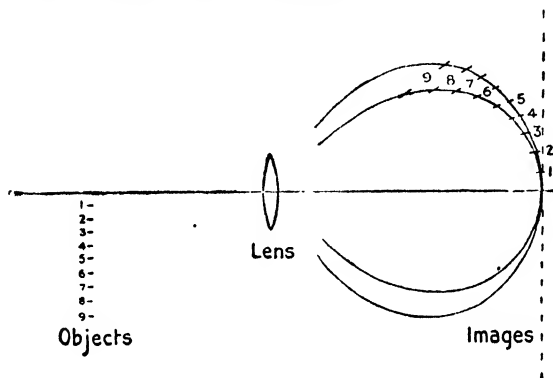
FIG. 45

nearer the lens being at right angles to the plane of the paper, and the one further from the lens being in the plane of the paper.

The defects are merely mentioned here as a reminder; they should be studied in the usual text-books.

§ 44. Depth of Focus

The image defects enumerated above may all be reduced or removed by polishing the surfaces slightly out of the spherical, by combining together several separate lenses made of special glasses, and by including stops or diaphragms at different positions in the compound lens.

There can, however, be no lens made which will produce, simultaneously, on a photographic plate *accurate* images of both near-by objects and distant mountains, and a photographer must decide which reproduced image he prefers to have accurately on his plate, or whether he prefers to effect

a material improvement in the definition of the one at the expense of a slight loss of sharpness of the other by placing the sensitive plate in some intermediate position.

Fortunately, it is not necessary to have absolutely accurate images, and in cinemas, on photographs, through telescopes, or through microscopes, the eye can recognize objects comfortably even if they are not absolutely 'pin sharp.' In normal photographs, for example, objects do not appear badly out of focus if the image produced by a point object is no larger in diameter than $\frac{1}{40}$ cm.,* and it is interesting to determine the range of object distances for which this will occur.
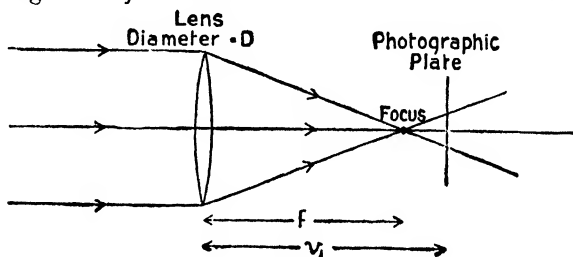
FIG. 46

Suppose the plate of a camera is set at a *position v_1* (Fig. 46) *such that the patch produced by parallel rays coming from a point object at infinity just appears as a point on the resulting photograph*, then the diameter of the patch must be not more than $\frac{1}{40}$ cm. Thus

$$\frac{1}{40} \text{ cm.} = \frac{D\left(v_1 - f\right)}{f}.$$

Now $$\frac{1}{v_1} + \frac{1}{u_1} = \frac{1}{f}, \quad \text{so} \quad \frac{1}{u_1} = \frac{v_1 - f}{f v_1}.$$

* For special work, such as enlarging, this value is much too great. An absolute limit is set to the resolution of the naked eye by the fact that the retina is made up of a collection of sensitive rods and cones separated from each other by 0·004 mm.; this limit actually corresponds to about 0·08 mm. at the nearest distance of distinct vision. Thus the conventional standard (0·25 mm.) for the sharpness of the image of a pin-point is somewhat large.

Thus
$$\tfrac{1}{40} \text{ cm.} = D\frac{v_1}{u_1}$$

and, for all normal cases in which u_1 is large,
$$\tfrac{1}{40} \text{ cm.} = D\frac{f}{u_1}.$$

If, therefore, the camera is set to focus accurately objects at a distance
$$u_1 = \frac{Df}{\tfrac{1}{40} \text{ cm.}},$$

it will give sufficient definition for objects ranging away to infinity.

Correspondingly, the nearer objects are to the lens than u_1, the larger becomes their patch on the photographic plate, and we can show to a first approximation that a camera set for a distance $\dfrac{Df}{(\tfrac{1}{40} \text{ cm.})}$ will have a depth of focus ranging from ∞ to one-half of this value. The following table constructed for a typical photographic lens of 10 cm. focal length shows how this depth of focus varies with the ratio $\dfrac{f}{D}$, which is the usual quantity lens-makers take to define the aperture of their lenses and which they usually mark on their adjustable diaphragms.

Table IV. Relation between Depth of Focus and Aperture of Lens for $f = 10$ Cm.

Aperture of lens $\dfrac{f}{D}$	1	3	4	6	8	16	32
Camera set for u_1	4,000 cm.	1,300 cm.	1,000 cm.	666 cm.	500 cm.	250 cm.	125 cm.
Range from ∞ to	2,000 cm.	650 cm.	500 cm.	333 cm.	250 cm.	125 cm.	62 cm.

The practical, important point to observe in this connection is that the depth of focus will be the greater, the smaller is D

and the smaller is f. It is for this reason that cameras are equipped with variable apertures so that when the illumination is good, or conditions are such that a long exposure is practicable, the aperture can be reduced to give greatly increased depth of focus. In the same way it is sometimes convenient to reduce the aperture of telescopes and, more often, of microscopes, in order to increase their depth of focus.*

§ 45. Magnifying Action of a Pin-hole

The direct relation between the depth of focus and the aperture of the lens is of much interest in connection with the behaviour of the eye. Whenever we try to see distinctly objects which are too close to be focused sharply on the retina, the iris automatically contracts, giving a smaller aperture and therefore increased depth of focus. Possibly part of the pain associated with defects of vision is due to the strain of the abnormally contracted pupils, and should, therefore, be similar to that normally encountered in strong sunlight.

Objects are usually brought very near to the eye in order to magnify them, that is, in order to increase the size of the image formed on the retina. When this image is large it covers more of the sensitive rods and cones of the eye so that more detail can be seen. Now it follows immediately from

* It should be noticed that the whole of the preceding treatment and table have been worked out on the assumption that the lens is perfect and that the light travels in geometrically straight lines. Actually we shall see, when discussing diffraction, that a perfect lens can never form a perfect point image. The demarcation lines between the central bright area and the first diffraction minima are shown by the dotted lines in Fig. 47, and it can be proved that its effective depth of focus is greater than that calculated above.
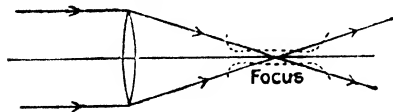
FIG. 47

In practice, however, the simple theory may be employed, since the errors in practical lenses usually make the depth much less than the theoretical value.

Fig. 48, *that the size of the image is proportional to the angle θ, so that it is usual to measure the apparent size of an object by the angle it subtends at the eye.*

A pin-hole in a paper or metal sheet held close to the eye
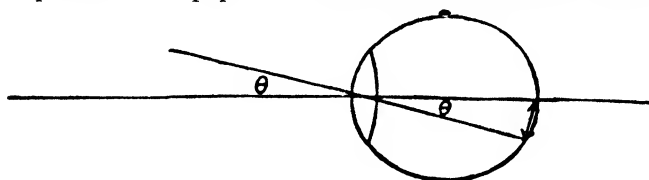
FIG. 48

provides the simplest device for magnification. Such a pin-hole increases the depth of focus, permitting clear images to be formed on the retina from objects which are placed close to the eye, and which, therefore, subtend a much larger angle at the eye than is normal.

§ 46. Magnification produced by a Lens held close to the Eye

There are two ways in which simple lenses may be used to magnify objects; the first is exemplified by the watchmaker's lens which is held close to the eye, while the second is exemplified by the simple magnifying glass used occasionally for reading.

The watchmaker's lens is virtually in contact with the eye lens and simply increases the power of the latter, making the wearer myopic or short-sighted, so that images of objects which are near to the eye can be formed on the retina.   The magnification produced is thus effectively the ratio of the nearest distance of distinct vision without the lens to that when the lens is used.

Clearly the magnification produced by this, or by similar devices, will depend on personal idiosyncrasies, such as minimum distance of distinct vision; in order, therefore, to compare the magnification produced by different arrangements, it is customary to employ in all such calculations a

standard distance of 25 cm. for the minimum distance of distinct vision.

We will now calculate the magnifying power of the single lens. Let f_e represent the focal length of the eye lens when the object is at the minimum of distinct vision u_e, and when the image is formed on the retina a distance v_e behind the eye lens.

Then
$$\frac{1}{u_e} + \frac{1}{v_e} = \frac{1}{f_e},$$

so that
$$\frac{1}{u_e} = \frac{1}{f_e} - \frac{1}{v_e}.$$

If now the lens f_l is placed in contact with the eye, the minimum distance of distinct vision will be reduced to u_l and the focal length of the combination is given by the equation:
$$\frac{1}{F} = \frac{1}{f_e} + \frac{1}{f_l}.$$

Thus, by analogy with the equation above:
$$\frac{1}{u_l} = \frac{1}{f_e} + \frac{1}{f_l} - \frac{1}{v_e}.$$
$$\therefore \frac{1}{u_l} = \frac{1}{f_l} + \frac{1}{u_e}.$$

Now the magnification produced by the lens is the ratio of the size of the image formed on the retina when the lens is present to its value when the lens is absent. Thus
$$\text{Magnification} = \frac{u_e}{u_l} = \frac{u_e}{f_l} + 1.$$

Notice that, in this calculation, we have found the magnification by taking the ratio of the angle subtended at the eye combination by the object at its nearest distance of distinct vision when the extra lens is and is not used respectively. Clearly a smaller numerical result for the magnification would have been found if the eye had been focused for infinity when the lens was used. If this had been the case, the ratio of the

two angles subtended—which is the magnification—would have been $\dfrac{u_e}{f_l}$, which is $\dfrac{25}{f_l}$ if f_l is expressed in cm.

In the other two conventions:

$$+ \frac{1}{u_e} = -\frac{1}{f_e} + \frac{1}{v_e}$$

and

$$\frac{1}{u} = -\frac{1}{f_e} - \frac{1}{f_l} + \frac{1}{v_e}$$

$$= -\frac{1}{f_l} + \frac{1}{u_e}$$

and magnification

$$= \frac{u_e}{u_l} = 1 - \frac{u_e}{l}.$$

In the convention in which lengths are positive if measured from the lens in the direction of motion of the light, u_e will be negative and f_l positive for a convex lens.

Finally, in the convention in which lengths are negative if measured from the lens in the direction of motion of the light, u_e will be positive but f_l negative for a convex lens.

Unless the lens is much more powerful than the eye, the observer will clearly place the object so that it appears as big as possible and so that the eye is working at its maximum power. If, however, the lens has a high magnification, it is most comfortable to place the object at the focus so that the eye itself is focused for infinity.

In practice it is difficult * to obtain single lenses with focal lengths less than one or two inches, because the shorter the focal length, the greater are the defects of the image, due to spherical and chromatic aberration. More complicated systems, such as the 'Wollaston doublet' and the 'Coddington lens,' reduce these defects considerably and render possible magnifications of 30 diameters or more. These systems have now, however, been completely superseded by the development of the compound microscope.

*D   * Such lenses have been made down to $f = \frac{1}{10}''$.

§ 47. The Compound Microscope

In the simplest form of compound microscope two lenses are used, the object glass—which forms an enlarged inverted image of the object, and the eyepiece—which again magnifies this enlarged image.
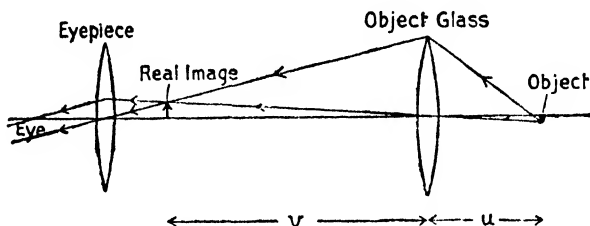
FIG. 49

It is clear from Fig. 49 that the real image is $-\dfrac{v}{u}$ times the size of the object, so that if f is the focal length of the eye lens (in cm.) the total magnification is (cf. previous paragraph):

$$-\frac{25}{f_1} \times \frac{v}{u} \quad \text{or} \quad -\left(\frac{25}{f_1} + 1\right) \times \frac{v}{u},$$

if the final image is formed respectively at infinity or at the nearest distance of distinct vision. As in the previous calculation, therefore, if the focal length of the eyepiece is small compared with 25 cm., very little increase in magnification will be produced by focusing the image at the near point.

Most compound microscopes are provided with a range of objectives so that the ratio $\dfrac{v}{u}$ can be altered at will. Since this is so, it is almost invariably undesirable to push the eyepiece magnification to its limit by adjusting its position to form the final image at the near point. It is best to adjust the eyepiece to produce its virtual image at infinity because every one can see objects at infinity, either with or without the aid of spectacles. The eyepiece will then always be in 'correct adjustment' for any observer, and it will be quite unnecessary for

any person to alter this eyepiece adjustment.   Moreover, the
eyepiece has been constructed to be most free from spherical
aberration and distortion when it is used to produce virtual
images at infinity.   A further point is that there is much less
strain on the eye in looking at an object at a great distance
than in looking at one which is situated at the near point.
Such strain is naturally small if continued for a few minutes
only, but may become important in extended work.

The ratio $\dfrac{v}{u}$ can be altered by increasing v and correspond-
ingly decreasing u, or by keeping v constant and altering both
u and the focal length of the objective.   As has been men-
tioned above, the latter is almost the invariable custom.
The objection to using a single objective and altering both v
and u is partly that the objective has been specially designed
to produce very low spherical and chromatic aberrations with
one particular value for u, and partly that changing the
magnification by a factor of 10 would mean very inconvenient
changes of u or v.

Two important mistakes are frequently made with micro-
scopes; the first with regard to the methods of drawing ray
diagrams, which are frequently drawn very unintelligently
without the student being aware he is making errors, and the
second with regard to the practical reasons for the use of
compound eyepieces.   These two mistakes will now be dis-
cussed in connection with telescopes, but it is important to
notice that the same arguments apply equally to micro-
scopes.

§ 48. The Simple Astronomical Telescope.   Ray Diagrams

In this telescope a long focus lens, called the object glass,
forms an inverted real image of the distant objects in, or
near, its focal plane.   This image is then magnified by a
powerful eyepiece held close to the eye.

The typical student's diagram shown in Fig. 50 is clearly
unsatisfactory, since the rays coming from the distant object

do not fall on the eyepiece at all, but *correspond to objects which are completely out of the field of view* of the instrument.
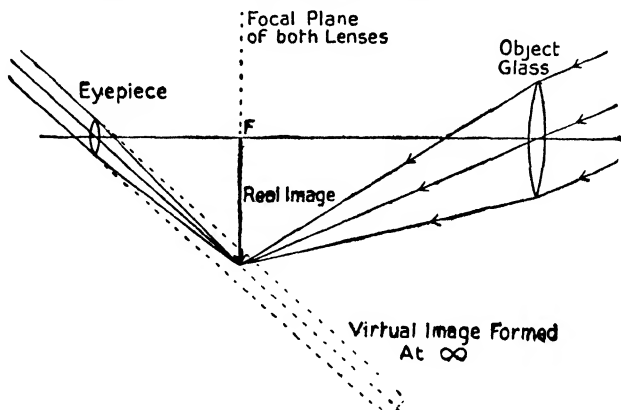
FIG. 50

Fig. 51 represents accurately the path of rays from an object which is just on the border of the field of view, for which only the rays from the part B of the object glass
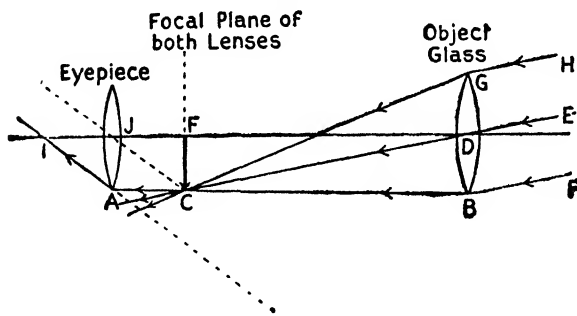
FIG. 51

reach the eyepiece. The following description of the method of drawing the diagram should be regarded merely as a practical recipe.

Draw the optic axis, the two lenses, and the focus (F).
Draw the line AB.
Mark the position of the real image FC.
Draw CDE.
Draw CG.
Draw GH and BF parallel to DE.
Dot in the constructional line JC.
Draw IA parallel to JC and extrapolate backwards.

§ 49. The Simple Astronomical Telescope.   Field of View

The field of view defined by the above diagram is quite small and will become progressively smaller as the distance between the lenses increases owing to longer and longer
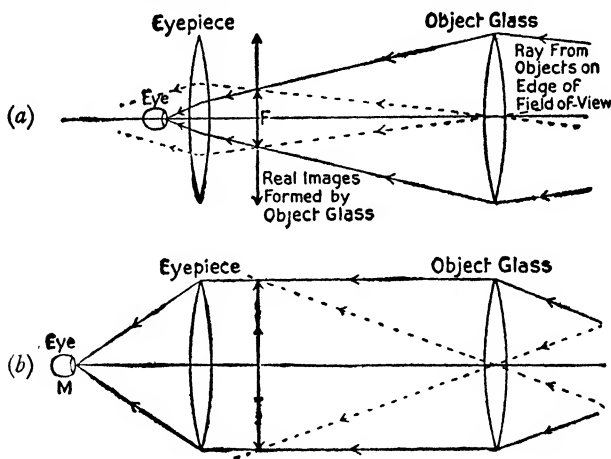
FIG. 52

focal length object glasses being used to increase the magnification. This minute field of view can be increased by increasing the diameter of the object glass, but there seems at first sight no advantage in increasing the aperture of the eye lens beyond that of the iris of the normal eye. This is, however, only correct if the eye is placed very close to the

eyepiece. If the eye is gradually withdrawn from the eyepiece the field of view will increase until it ultimately fills the whole of the eyepiece.

Fig. 52 (a) shows the field of view when the eye is placed close to the eyepiece, while Fig. 52 (b) shows that if the eye is removed to the point M it will receive rays of light from a much wider field of view, the rays coming from different parts of the object glass.

In practical instruments it is usual to fix a metal or ebonite ring to prevent observers from moving their eye from this position of maximum field of view. This ring is called the eye ring of the instrument.

In the best position the point M should lie on the centre of rotation of the eyeball, because when a large field of view is presented to the eye, the latter rotates until the image is formed on the most sensitive spot of the eye—the *fovea centralis*. The observer is, however, aware of objects which are remote from this carefully studied portion, and can immediately concentrate on any other area which has attracted his attention.

§ 50. The Astronomical Telescope. Compound Eyepiece

The field of view when the eye is close to the eyepiece can be enlarged by as much as 5 or 10 diameters by the addition of an-
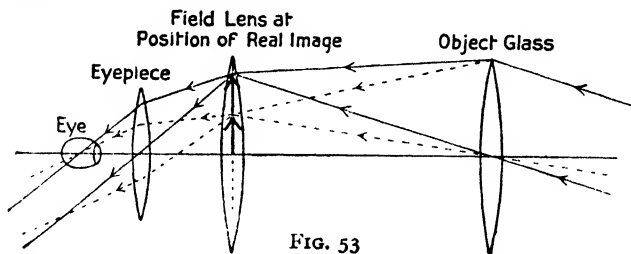
Fig. 53

other lens placed in the focal plane of the object glass and eyepiece (Fig. 53). This lens should, for preference, have nearly the same focal length as the eyepiece, and should have as

large a diameter as possible. Its presence cannot in the least alter the correct setting of the distance between object glass and eyepiece, because the real image is formed on the field lens itself, so that u and therefore v are each zero. The field lens, therefore, merely bends the rays so that they enter the eyepiece, but does not affect their divergence.

The first advantage of the simple field lens is thus to bring the eye ring closer to the eye and to increase still more the field of view of the instrument; it has, however, a further slight advantage since it serves to increase the brightness of images at the edge of the field of view.

§ 51. The Astronomical Telescope.   Huygens and Ramsden Eyepieces*

The simple field lens that we have just discussed is very easy to understand theoretically, since its inclusion does not affect the position of the real image or the magnification of the telescope. Several more complicated compound eye-pieces have been produced to suit the particular needs of individual workers by varying the positions and the focal lengths of the field lens and eyepieces. It is, of course, clear that any field lens placed in almost any position in the compound eyepiece must increase to some extent the field of view when the eye is close to the eyepiece. In addition, it will also affect the distortion and the spherical and chromatic aberrations of the instrument. If the field lens is not at the focus of the eyepiece, the magnification of the telescope will naturally be altered, but can readily be calculated by regarding the eyepiece and field lens as a single thick lens of focal length F.

* Messrs. Griffin & Tatlock supply a very convenient model telescope and microscope which shows immediately the practical effect of changing from a Galilean to an astronomical, and from a simple field lens eyepiece to a Huygens or a Ramsden. Similarly, a simple microscope objective or a Wollaston objective can be used.

The magnification is then

$$= \frac{(f \text{ of object glass})}{F},$$

where F is given by the equation:

$$\frac{1}{F} = \frac{1}{f \text{ eyepiece}} + \frac{1}{f \text{ field lens}} - \frac{a}{f \text{ eye } f \text{ field}}.$$

The following two compound eyepieces are most common:

THE RAMSDEN EYEPIECE

This eyepiece (Fig. 54) has exceedingly low spherical aberration, and has a very flat field of view, but is not completely achromatic. It is very useful with cross wires or eyepiece scales, which are mounted at the position of the real image.
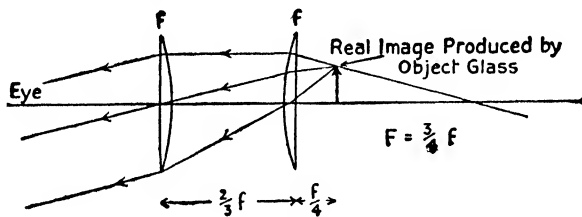
FIG. 54

THE HUYGENS EYEPIECE

This eyepiece has not such a large magnification as the Ramsden eyepiece. It has a slightly smaller field of view if the two field lenses have the same diameters, but, since the Huygens field lens is of longer focal length than the Ramsden field lens it usually has a larger diameter and so gives the same, or possibly a larger field of view than is obtained with a Ramsden of the same cost.

The Huygens eyepiece is completely achromatic, and has low spherical aberration—though not as low as the Ramsden. Although cross wires at f_e from the eye lens can be used, a scale in the eyepiece is generally unsatisfactory because

considerable distortion occurs at this point so that equal distances on the scale would not correspond to equal angles of the incident rays.
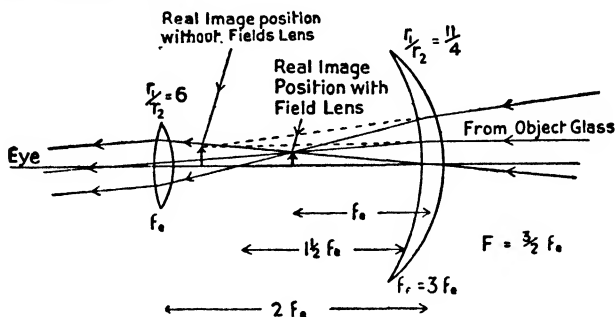
FIG. 55

The actual figures quoted in Fig. 55 for the Huygens eyepiece form the most perfect example of this type, but a very considerable latitude of radii of curvature and even of focal length is permissible, since the term 'Huygens eyepiece' is often used in a rather general way.

§ 52. Telescopes and Achromatism

If the objective of a telescope were perfectly achromatic,* it would form one single image of a point source independent of the colour of the light used. If, in addition, the eyepiece were also perfectly achromatic, the virtual image formed would be precisely the same for all colours.

It is, however, quite unnecessary for the eyepiece of a telescope to be achromatic, if it is composed of a single thin lens. This will be seen from Fig. 56, which shows that the separate coloured virtual images of the single image

* It is impossible to construct perfectly achromatic lenses. Normal achromatic lenses are simple doublets having identical focal lengths for one wave-length in the red and for one wave-length in the blue. Achromatic triplets would have the same focal length for three different wave-lengths, but no lens can have exactly the same focal length for *all* colours, although certain makes show departures which are only just measurable.

formed by the achromatic objective all lie completely
behind each other and so do not show coloured edges.

Virtual Image for
Violet Rays

Virtual
Image for
Red Rays

Achromatic
Object Glass

Eyepiece

Eye

Single Real Image
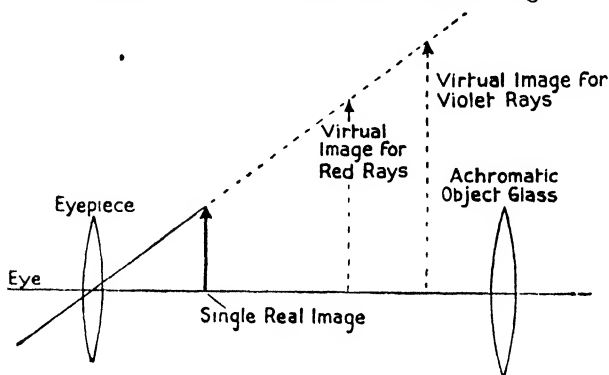
FIG. 56

Fig. 56 should be contrasted with Fig. 57, which holds good
if the object glass is not achromatic and which is nearly the
same whether the eyepiece is or is not achromatic.

Eyepiece

Non
Achromatic
Object Glass

Eye

Real
Red
Image

Real Blue
Image

Virtual Red
Image

Virtual
Blue
Image
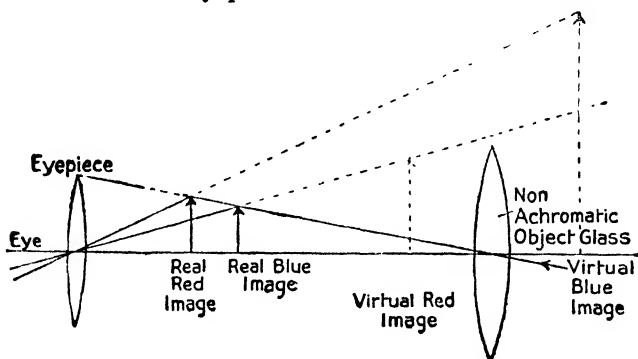
FIG. 57

Naturally, if the eyepiece is compound, the same type of
argument applies, so that the field lens should, in general, be
achromatic. This is fortunately quite unnecessary with the
Huygens eyepiece, and is usually unnecessary with a Ramsden

eyepiece—if the two lenses are made of the same glass; since detailed theory shows that the *magnification* of these systems is nearly independent of the colour.

§ 53. Terrestrial Telescopes.   Galilean

The telescopes which have so far been discussed are unsuited for terrestrial observation, because of the inversion of the image.   The first terrestrial telescope forming an erect image was improved by Galileo and is described by Fig. 58.
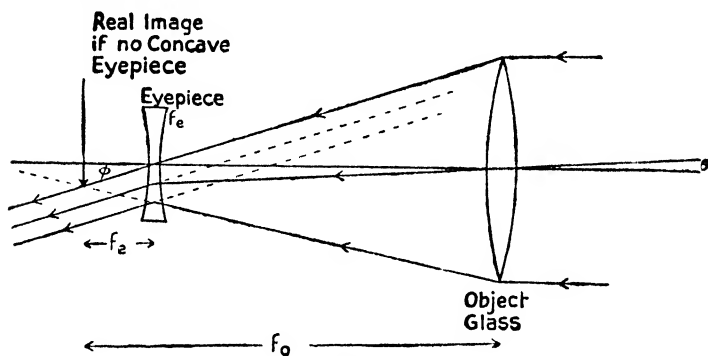
FIG. 58

The magnification is the ratio $\dfrac{\phi}{\theta}$.

But $\phi = \dfrac{\text{real image size}}{f_e}$ and $\theta = \dfrac{\text{real image size}}{f_0}$,

so $\text{magnification} = \dfrac{f_0}{f_e}$.

The above calculation of the magnification is easy only if the diagram has been drawn correctly with one of the rays from the object glass passing through the centre of the eyepiece; it would have been more difficult if the diagram had been as in Fig. 59.

A comparison of Figs. 58 and 59 with the corresponding

ray figure for an astronomical telescope (Fig. 52) shows that with a Galilean eyepiece the rays coming from different parts of the field of view are diverging, so that there is no eye ring, and the field of view decreases as the eye is withdrawn from the eyepiece.
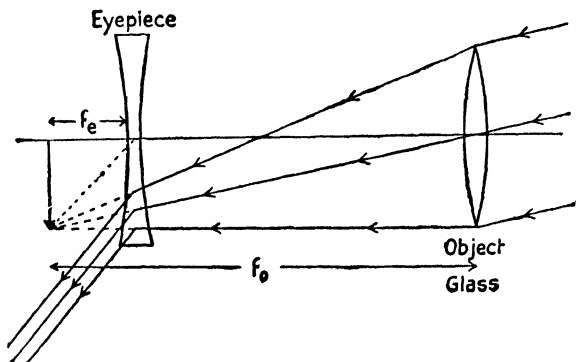
FIG. 59

The great disadvantage of the Galilean telescope is, therefore, that the maximum field of view is very small—the emergent magnified field having a total angular diameter of only $\left(\dfrac{\text{diameter of object glass}}{f_0 - f_e}\right)$, so that the larger the magnification the smaller is the field of view.

§ 54. Terrestrial Telescopes.   Erecting Lenses and Prisms

On account of the small field of view of the Galilean telescope compared with that of the astronomical telescope, it is natural to attempt to modify the latter to produce erect images, and so to use the field lens assembly or the simple eye-ring properties of the converging eyepiece.

There are two ways in which the inverted image of the normal astronomical telescope can be made erect.   The first

of these is by the inclusion of an erecting lens, and is described sufficiently by the simple diagram shown in Fig. 60.
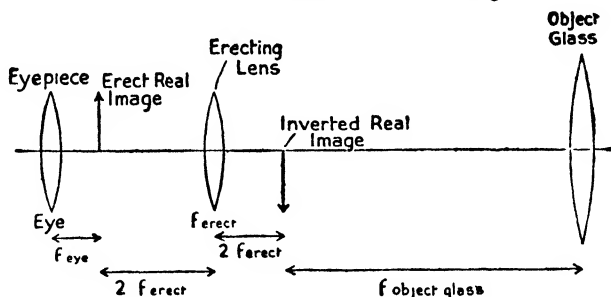
FIG. 60

The second, and most common, method is to erect the image by means of two right-angle prisms. Since it is very easy to get incorrect or, more commonly, incomplete ideas on this matter, the items are worth considering in detail:

(1) Fig. 61 shows that a 90°, 45°, 45° prism reflects back through exactly 180° all rays incident on its hypotenuse. This will occur even though the rays are not incident exactly normal to the surface.
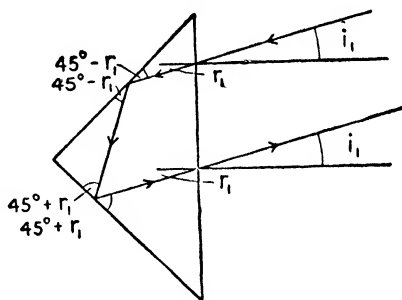
(2) The prism should be made of glass with a high refractive index,

FIG. 61

otherwise some of the rays will not have a sufficiently large angle of incidence at the glass-air layer to be totally reflected, and much light-energy will be lost.

(3) Fig. 62 shows how the prism inverts the image which is being formed by the lens. It is confusing to trace the actual path of the rays through the prism, and all that is necessary is to note that each of the rays has been turned through

exactly 90° and that the one which is incident furthest from the centre is reflected furthest from the centre and is inverted.
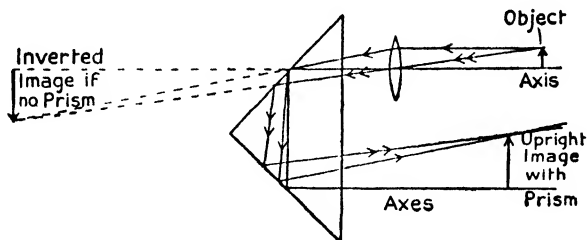
FIG. 62

(4) The prisms are *not* arranged as shown in Fig. 63. If they were, the first would erect the image and the second invert it again.
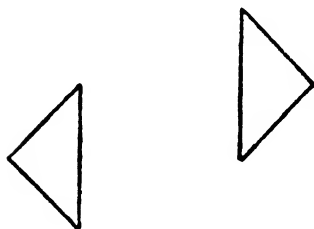
FIG. 63

IMPORTANT

(5) The prisms are arranged to erect first in the plane of the paper and then in the plane at right angles to the paper, as shown in Fig. 64.

(6) The advantage of a prism monocular is essentially that it possesses the field of view characteristic of an astronomical telescope with a field lens, and that this field of view is markedly greater than that of a similar Galilean telescope.

(7) The prisms naturally shorten the length of the instrument compared with that of the astronomical telescope which could have been constructed from the same lenses. This is generally an unimportant matter compared with the importance of the increased field of view. In practice, as a matter of fact, prism binoculars are frequently almost as long as, or even longer than, Galilean glasses, presumably because the latter often use more powerful lenses.

(8) If prism binoculars are used, the prisms enable the two object glasses to be separated more than are normal eyes.
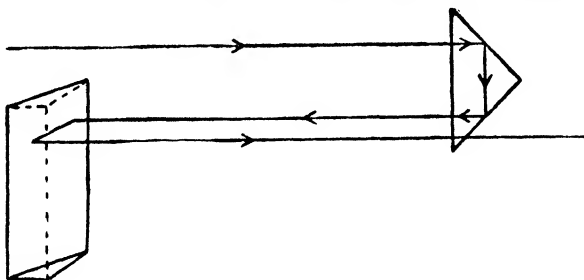
Fig. 64

They therefore exhibit a pronounced stereoscopic effect, and show depth even more clearly than normal eyes.

§ 55. The Purchase and Use of Telescopes

The most important particulars to be considered in the purchase of a telescope system are (apart from the cost) the magnification, the field of view, the light-gathering capacity, and the sharpness and lack of colour or distortion of the final image.

For terrestrial work it is unusual to have a magnification greater than 8 or 16 diameters, because, unless the instrument has been exceedingly carefully designed and, therefore, is very expensive, the larger the magnification the smaller is the field of view. Toy telescopes costing two or three shillings can be made with very large magnifications indeed, but this is at the expense of the field of view, and any instrument combining high magnification with large field of view will be expensive.

If required for use at night, the most important point concerning the instrument is its light-gathering capacity. The brightness of the image on the retina will be greater, the greater is the ratio

$$\frac{\text{lens area}}{\text{eye-pupil area}}$$

(providing that all the light collected by the field lens enters the eye*). This brightness will be less, the greater is the ratio

$$\frac{\text{retinal image area with lens}}{\text{retinal image area without lens}},$$

which is, of course, m^2. Thus the brightness of the image on the retina is proportional to

$$\frac{\text{lens area}}{\text{eye-pupil area}} \times \frac{1}{m^2}.$$

Now it is an interesting fact that if this is calculated for any particular system, it is always 1 or less than 1. (Refer to footnote below). Thus it is impossible to increase the brightness of the retinal image of terrestrial objects by the use of lens systems, and in practice—owing to the light-absorption of the many lenses employed—telescopic systems reduce considerably the intrinsic brightness. The simple Galilean telescope is very efficient in this respect and so is better adapted than prism binoculars for work under conditions of poor illumination.

It is possible to increase very greatly the brightness of images of distant stars by increasing the area of the object glass, because *with stars, and with stars alone*, the image produced is still a point, and is so much smaller than the normal diffraction haloes surrounding the image, that an increase in magnification produces no effective increase in size of the image. Thus, with stars a larger telescope objective not only collects more light, but actually concentrates the light into a smaller area, because, as we shall see later, the size of the diffraction patch produced by a lens decreases as the diameter is increased.

It is very desirable that optical instruments such as telescopes and microscopes be adjusted with the final, virtual,

---

* This clause is important because it shows that there can be no advantage in brightness in increasing the size of the objective of a telescope beyond a certain limit. This limit must clearly be dependent on the magnification, and if we take ⅕ inch for the size of the eye pupil, the limiting diameter of a telescope objective is $\frac{m}{5}$ inch.

image at infinity, so that they may be used with the un-accommodated eye. There is then no strain on long continued observation, because every person can conveniently focus parallel rays—or wears glasses to enable him to do so. An additional reason is that the lenses of the instrument are corrected for their various image defects on the assumption that they will be used in this way. Inexperienced users accommodate for near vision, merely to increase slightly the magnification, and do not notice that the definition is then so much poorer that they can see little, if any, more detail than was possible with the unaccommodated system.

A convenient way to test if the emergent rays are really parallel and coming from a virtual image at infinity is to place another converging lens in front of the eyepiece, and to see if a real image is formed in the focal plane of the lens.

Experimental work on the magnification of telescopes can be readily performed by the usual method of observing an object directly with one eye while the other is viewing the same object through the telescope. This method naturally fails completely when the telescope is focused for the nearest distance of distinct vision, and for such investigations an accessory is desirable. This accessory has two lenses—of 10·0 and 7·5 cm. focal length respectively—mounted exactly 10 cm. from a ground-glass scale. The first of these lenses brings parallel rays to a focus on the ground glass, and thus represents an unaccommodated eye and retina, while the second represents an accommodated eye and focuses objects at the nearest distance of distinct vision—30 cm. The distance between the images of two distant lights is then found, using first direct light, then light which has passed through the telescope focused for infinity, and finally, light which comes from the telescope when focused for the nearest distance of distinct vision. The ratio of the separation between the images gives directly the magnification in the two positions.

CHAPTER V

THE NATURE AND POLARIZATION OF LIGHT

THE whole subject of the polarization of light is complicated by the fact that students generally have no real physical picture of the nature of light, and rely too much on vaguely understood statements of the type: 'Light is an electromagnetic disturbance consisting of transverse vibrations with the electric vector perpendicular to the direction of propagation of the wave,' or 'since the amplitude of a spherical wave is inversely proportional to the distance from the origin of the wave, the energy varies inversely as the square of that distance.' The following simplified picture of the mechanism of light propagation may remove some difficulties.

§ 56. Tubes of Electric Force and the Propagation of Light

If an electron is moving forward with a constant velocity (v), all its lines of force will also be moving forward uniformly with the same velocity. Suppose now that the electron is brought completely to rest in a very short time which we will call τ, then the lines of force may, or may not, also be brought to rest instantaneously. Let us assume that they are not, but that instead, the information that the electron has stopped is sent along the line of force with a definite velocity, which we may, for convenience, call c. This is not such an absurd thing to assume as it at first appears, because lines of force

106

behave in many respects like stretched elastics—and we do know that such elastic strings can transmit waves with finite velocities.   Moreover, if the lines do, after all, stop instantaneously, our theory will be valid when c is given the value of infinity. Actually, of course, it is possible with more advanced work to calculate the value of the velocity c and to show that it is $\dfrac{1}{\sqrt{\mu k}}$ and is therefore 3×10^{10} cm. a second, which happens to be exactly the same as the value found experimentally for the velocity of wireless, heat, and light waves; it is largely this agreement which makes us say that light waves

FIG. 65 (a)

are really due to the 'kinks' in lines of force coming from an electron which is suddenly retarded.

Fig. 65 (a) represents the appearance of some of the lines of force emitted by the electron at a time t after its stoppage, and we will assume that τ (the short time in which it was retarded) is very small indeed compared with t.   The lines of force below the points AB are radiating from the actual position of the electron E, while the lines of force above CD are radiating from the position where the electron would have been at this instant if it had not been stopped.   They are, therefore, radiating from a point F, which is a distance vt further on (or $v(t + \tau)$, it does not matter which).   Between
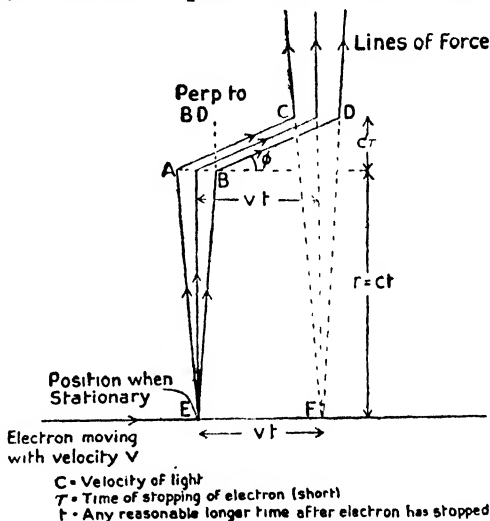
AB and CD the lines of force must be bent as shown, and this bend in the lines is moving outwards with a velocity c.

Now the force on an electron in a field is (*the strength of the field*) × (*the charge on the electron*) which is (*the number of lines per sq. cm*) × (*the charge*). Over the portion AB the concentration or density of the lines of force is $\frac{e_1}{kr^2}$, and therefore, over the plane perpendicular to AC and BD, the concentration is $\frac{e_1}{kr^2 \sin \phi}$ so that an electron e_2 placed in this region experiences a force of $\frac{e_1 e_2}{kr^2 \sin \phi}$ in the direction shown here: Fig. 65 (*b*).
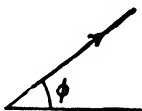
This force can be split up into two components, **the one** along the radius EB and the other at right angles **to it.** These components are:

(1) Radial $\quad\quad \left(\frac{e_1 e_2}{kr^2 \sin \phi}\right) \times \sin \phi.$

(2) Perpendicular $\left(\frac{e_1 e_2}{kr^2 \sin \phi}\right) \cos \phi$

$$= \frac{e_1 e_2}{kr^2} \cot \phi = \frac{e_1 e_2}{kr^2} \left(\frac{vt}{c\tau}\right),$$

since $\cot \phi$ is clearly $\left(\frac{vt}{c\tau}\right).$

It follows, therefore, that the electron experiences an unchanged steady repulsion in the direction EB, but that, for the very short time the pulse or kink is present, it also experiences a perpendicular component of force of

$$\left(\frac{e_1 e_2}{kr^2}\right) \left(\frac{vt}{c\tau}\right)$$

which is $\quad \left(\frac{e_1 e_2}{kr^2}\right) \frac{v}{c\tau} \times \left(\frac{r}{c}\right) = \frac{e_1 e_2}{krc^2} \frac{v}{\tau}.$

Thus, whenever an electron suddenly either begins to move, or is brought to rest, others, in a plane at right angles to its direction of motion, will also experience forces which are perpendicular to the radius joining them, that is, perpendicular to the direction of propagation of the electric wave.

Further, we may notice that the ratio $\frac{v}{\tau}$ is the acceleration of the primary electron, and so, whenever an electron is accelerated or retarded, others will experience a force of

$$\frac{e_1 e_2}{kc^2 r} \times \text{(acc.)}.$$

This force describes the *amplitude of the wave* and *varies inversely with the distance* of the emitting electron.

The *electrical energy Per Unit Volume of the wave must vary inversely as the square of the distance* from the primary electron, because an electrostatics theorem shows that it is

$$\frac{k(\text{electric field})^2}{8\pi},$$

which is clearly

$$\frac{k}{8\pi}\left(\frac{e_1 \times \text{acc.}}{kc^2 r}\right)^2.$$

It is important to note that we have so far been considering only one half of the total phenomenon involved in the electromagnetic wave. The perpendicular component of the electric field $\frac{e_1 f}{kc^2 r}$ is moving perpendicularly to itself with a velocity c. It is, therefore, producing a magnetic field moving with the velocity c, and this field is at right angles to the direction of propagation and to the direction of the electric field. The combined electric and magnetic fields constitute the complete electromagnetic wave.*

* It is easy to write down the magnitude of this magnetic field and to prove that the energy in the electric field is exactly equal to that in the magnetic field. This is necessary, for example, when one is using the above theory in order to calculate the magnitude of the pressure which light exerts on an absorbing surface in virtue of the

§ 57. Electromagnetic Induction and the Propagation of Light

An alternative method of treatment has considerable interest.

According to Ampère's theorem on the magnetic field produced by current elements, the magnetic field produced at right angles to a current element ids is

$$H = \frac{ids}{r^2}.$$

Now the current i must be

$$i = nv(e_{emu}),$$

if n denotes the number of electrons per cm. length of the wire, v their velocity, and (e_{emu}) the charge in electromagnetic units on each electron.

Thus $$H = nv\left(\frac{e_{emu}}{r^2}\right)ds$$

and—since nds is the number of electrons in the current element —the field produced at the point by one electron moving with a velocity v will be

$$H = v\frac{(e_{emu})}{r^2}.$$

Now this field is radial round the electron, and is coming perpendicularly from the paper; so if we were to place a very small rectangular loop of copper wire of sides dh and dr, as shown in Fig. 66 (a), at the distance r from the electron the total number of lines of force inside it would be
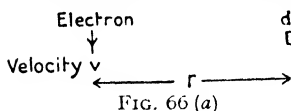
Electron
Velocity v
Fig. 66 (a)

$$\frac{v(e_{emu})}{r^2}dhdr,$$

and the total number of lines of induction

$$\mu\left(\frac{v(e_{emu})}{r^2}\right)dhdr.$$

action of the magnetic field on the secondary electrons set in motion by the electric field.

The magnetic field is also necessary when this theory is being extended to determine the total quantity of energy radiated by an accelerated electron as is commonly explained in the usual text-books on X-rays.

If we now consider the electron to be changing its velocity with a constant acceleration, there must be a change in the number of lines of induction in our tiny wire loop. In consequence the E.M.F. set up will be

$$\frac{dN}{dt} = \frac{\mu \frac{dv}{dt}\,(e_{emu})}{r^2}\,dhdr.$$

Now consider a wire loop of the shape shown in Fig. 66 (b) stretching away to infinity. The total E.M.F. set up will be

$$\mu(\text{acc.})\,(e_{emu})dh \int \frac{dr}{r^2} = \mu(\text{acc.})\,(e_{emu})\frac{dh}{r}.$$

This E.M.F. will be perpendicular to the lines of magnetic induction and to the direction of their motion; it will therefore be parallel to the direction of the acceleration of the original electron, and will be in the direction of dh. If the wire loop were absent

Electron
↓
Velocity v

dh ▭

←——— r ———→

FIG. 66 (b)

we should therefore observe an electric field produced whose magnitude would be

$$X_{emu} = \frac{\text{potential diff.}}{dh} = \mu(\text{acc.})\frac{(e_{emu})}{r}.$$

Thus once again we see that there is produced an electric field which varies in intensity according to $\frac{1}{r}$, and whose energy

$$\frac{kX^2_{esu}}{8\pi}, \text{ must be } \frac{k}{8\pi}\frac{X^2_{esu}}{X^2_{emu}}\left\{\mu(\text{acc.})\left(\frac{e_{emu}}{r}\right)\right\}^2,$$

varies inversely as the square of the distance from the primary electron.

Now since the force on an electron must be the same in e.m. as in e.s. units

$$X_{emu}\left(e_{2em}\right) = \mu_{e.m.}\left(\text{acc.}\right)\frac{e_{em}}{r}\,e_{2emu}$$

$$\text{is also} = \frac{e_{es}}{k_{es}}\frac{e_{2es}}{c^2}\frac{1}{r} \times \left(\text{acc.}\right)$$

$$\therefore c^2 = \frac{X_{em}}{X_{es}} \times \frac{e_{es}}{e_{em}} \times \frac{1}{\mu_{em}\,k_{es}}$$

$$= \frac{1}{k_{em}} \times k_{es} \times \frac{1}{\mu_{em}\,k_{es}} = \frac{1}{\mu_{em}\,k_{em}}.$$
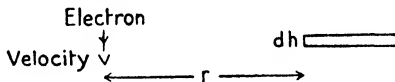
§ 58. Light Intensity at an Angle with the Electron Acceleration

If the direction of the acceleration (f) of the primary electron is at an angle θ with the direction of the emission of the light rays, the primary acceleration can be resolved into two components, $\rightarrow f \sin \theta$ and $\uparrow f \cos \theta$ (Fig. 67).
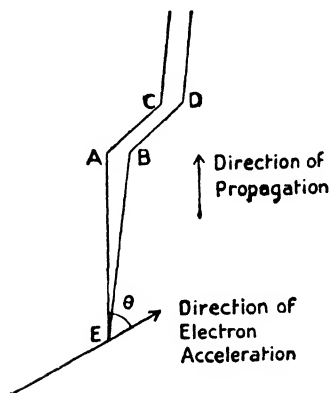
The amplitude due to the first of these will then be

$$\frac{e_1 e_2}{kc^2 r} f \sin \theta,$$

while that due to the second can easily be shown to be zero.

Once again, therefore, we have a force on the second electron which is perpendicular to the direction of propagation and which is inversely proportional to r.

↑ Direction of
  Propagation

Direction of
Electron
Acceleration

FIG. 67

§ 59. Emissions from Electrons executing Simple Harmonic Vibrations

In the equation we have just found, the amplitude of the waves, or the force on the secondary electron, is proportional to the acceleration f of the primary electron. This is of particular interest when the primary electron is moving in a simple harmonic motion represented by the equation:

$$x = x_0 \cos pt,$$

because then the acceleration f is $-p^2 x$, so that the force on the secondary electron is

$$\frac{e_1 e_2 \sin \theta}{kc^2 r} p^2 x_0 \cos pt,$$

and is, therefore, also a simple harmonic of the same frequency.

Thus we see that if the primary electron oscillates with a simple harmonic motion the force on a secondary electron

will also oscillate with the same period in a direction at right angles to the radius joining the two electrons.   The amplitude of the secondary oscillating force will vary as $(\sin \theta)$.

A wave of this character sent off by a radiating electron will be called a *plane polarized wave* because, for any ray, the oscillations of both the primary and the secondary electrons are in one single plane.

§ 60. Circularly and Elliptically Polarized Light

Many of the electron oscillations occurring in nature are those in which the electron is moving in one plane along a circular or elliptic path round a central positive charge. Now any elliptic vibration can be regarded as being made up of two simple harmonic vibrations of the same frequency p along axes which are at right angles to each other.   Thus an elliptic vibration can be represented by the equations:

$$x = x_0 \cos (pt),$$
$$y = y_0 \cos (pt + \delta).$$

If you do not believe this, you can easily remove t from these two equations and find the locus of the point, i.e. the equation to the curve on which the electron must be at any and every time you may consider.

This calculation can be carried out as follows:

$$y = y_0 \cos (pt + \delta)$$
$$y = y_0 \cos (pt) \cos \delta - y_0 \sin (pt) \sin \delta$$

and
$$x = x_0 \cos pt.$$

Thus:

$$y = y_0 \left(\frac{x}{x_0}\right) \cos \delta - y_0 \sin (pt) \sin \delta$$

$$\left(y - \frac{y_0 x}{x_0} \cos \delta\right)^2 = y_0^2 \sin^2\delta \, (1 - \cos^2 pt)$$

$$= y_0^2 \sin^2\delta \left(1 - \frac{x^2}{x_0^2}\right)$$

and so
$$\left(\frac{y}{y_0}\right)^2 - \frac{2xy}{x_0 y_0} \cos \delta + \left(\frac{x}{x_0}\right)^2 = \sin^2 \delta$$

E

which is the equation of an ellipse on axes which are askew and which are not x and y. Fig. 68 shows the graph of this equation and therefore represents the path of the electron that we have so far been considering. If we now consider the behaviour of *another* electron situated at any point Z (Fig. 68), on the axis through O perpendicular to the paper, then this electron will clearly oscillate parallel to OX under the action of the forces (see equation 1, p. 113), produced by the X component of the motion of the first electron. It will also oscillate parallel to (OY) under the action of forces due to the y component of the motion of the primary electron. The electron at Z will, therefore, move in the same type of ellipse as did the primary electron in the plane XOY, and the wave travelling in the direction OZ will be called an *elliptically polarized wave*.
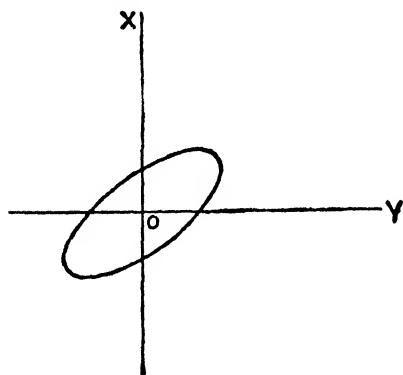
FIG. 68

If the primary electron vibration had happened to have $x_0 = y_0$ and had had $\delta = \frac{\pi}{2}$, then the locus of the electron would have been $y^2 + x^2 = x_0^2$, which is a circle. The secondary electron at Z would then also have experienced forces tending to make it rotate in a similar circle, and the wave travelling in the direction OZ would be called a *circularly polarized wave*.

It is particularly interesting to refer again to Fig. 68 and to consider the forces acting on an electron which is some distance from the first electron and which is in the direction OX. This electron will clearly be subjected to forces which

are parallel to OY on account of the oscillation of the primary electron in the direction OY. The corresponding forces due to the primary oscillation in the OX direction will not affect the electron at X, because the force is $\dfrac{e_1 e_2 \sin \theta}{kc^2 r} p^2 x_0 \cos pt$ (see page 112), and, at the point X, $\sin \theta = 0$ for the oscillation in the OX direction. Thus the light emitted along rays which are in the plane XY, by an electron which is vibrating in an ellipse in the plane XY, will be plane polarized with its electric vector in the plane XY.

§ 61. The Nature of Ordinary Unpolarized Light

According to the simple physical picture presented above, the light emitted by a single vibrating electron system should be plane, circularly, or elliptically polarized, so that we have, at present, given no explanation of normal unpolarized light.

Now bodies which are emitting white light as a result of internal temperature agitation can be regarded as being made up of innumerable electron systems oscillating with every conceivable frequency. Moreover all those electron systems which have identically the same frequency and so have the same colour, have the complete range of phase differences between the X and Y components of the electric force, so that in these directions the total electric force will be the sums of the components due to each vibrating electron in turn, as follows:

In X direction total force
$$= x_{01} \cos (pt + \delta_1) + x_{02} \cos (pt + \delta_2) + \text{etc.}$$
In Y direction total force
$$= y_{01} \cos (pt + \delta_1) + y_{02} \cos (pt + \delta_2) + \text{etc.}$$

These separate components when added together naturally have a resultant with a total amplitude which will be statistically constant.

Although the resultant amplitude will be statistically constant, the phase differences between the X and the Y

resultants will be continually varying as the individual electrons stop and restart at each intra-molecular collision. Thus this temperature radiation is neither plane, circularly, nor elliptically polarized, but is, in fact, instantaneously varying between these types in a completely random manner.

§ 62. Polarization by Reflection

We will now discuss methods by which it is possible to isolate either the *x* or *y* resultants of unpolarized monochromatic light.

Incident Ray

X Component perp. to Paper
Y Component

Reflecting Surface
Plane perp. to Paper
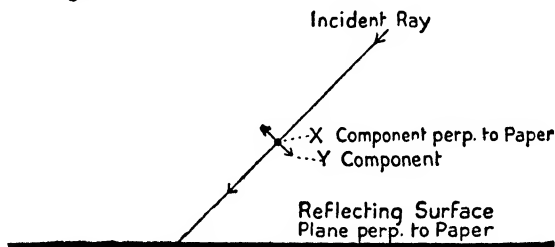
FIG. 69

Consider a ray of unpolarized light (that is, light with irregular and varying phase differences between its *x* and y resultants) falling at an angle θ on a glass or metal reflecting surface. The electrons in the reflecting surface will be forced by the *x* component of the electric force to vibrate up and down, i.e. at right angles to the paper in Fig. 69. They thus vibrate in the plane of the reflecting surface. Correspondingly, the *y* component will make the electrons vibrate in the direction shown—which is, to some extent, in and out of the plane of the reflecting surface.

There seems to be an essential dissimilarity between an electron vibration in the plane of the surface and one which is largely in and out of it. It is, therefore, not surprising that the two vibrations are not treated similarly by the reflecting surface, and that the one is markedly more strongly reflected than the other.

Experiment shows that with substances such as glass, with refractive indices between 1·4 and 1·5, the X component *alone*, with its electric field in the plane of the reflecting surface, is reflected if the angle of incidence is about $\tan^{-1} \mu$. The whole of the Y component—the in-and-out one—is then transmitted into the glass together with, of course, the remainder of the X component.

The plane containing the incident ray, the normal, and the reflected ray, is called the plane of polarization of the light, and so, according to this conventional nomenclature, the
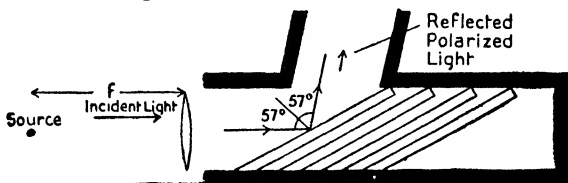
FIG. 70

electric vector is perpendicular to the plane of polarization of the light.

The broad principles of the experimental technique for obtaining light plane polarized by reflection, can be seen from Fig. 70. The beam of parallel rays falls on to a number of glass plates placed at the polarizing angle (which is about 57° for glass). The transmitted light, and any stray, scattered light, is completely absorbed by the black lining of the container holding the glass plates. The advantage of using a large number of plates is that the intensity of the reflected polarized light is considerably enhanced.

§ 63. Polarization by Double Refraction

The reduction in the velocity of light when it is traversing any transparent medium such as glass or crystal is directly consequent on the existence in the material of bound electrons, and more advanced theory can connect the refractive index of any particular wave length with the number and natural free period of these electrons. This knowledge permits

us to obtain a general physical explanation of the polarization effects observed in some crystals.

The essential difference between crystalline and amorphous bodies is that, in crystals, the atoms or molecules are arranged in a definite, orderly, and symmetrical fashion.

There are many different ways of packing molecules together in orderly arrangements, and natural or artificial examples of every one of these modes can be obtained. The simplest is probably that in which the molecules are packed at the

FIG. 71

All distances equal
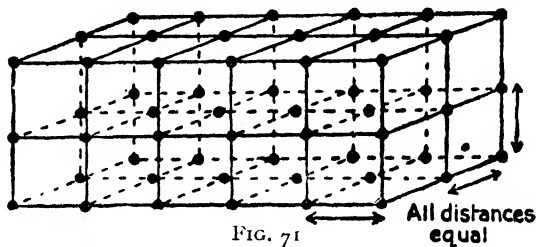
corners of cubes as in Fig. 71; here, the distances in the x, y, and z directions are all equal.

The somewhat less simple arrangement that occurs with crystals which crystallize in the tetragonal system is identical with the simple cubic structure with the exception that the distances between the atomic units in one direction (say the Z direction) are appreciably different from those in the other two. This is exemplified by Fig. 72, where the dots represent the fundamental units of the structure, whether these units are atoms, compound ions, molecules, or even molecular aggregates.

In order to form a clear idea of the subject, it is worth repeating that crystals can be found in nature to illustrate each in turn of all the possible internal orderly arrangements which can possibly occur. All the crystals showing a simple cubic arrangement of their constituents behave, to the propagation of light, in the same way as glass or other isotropic bodies, but all those crystals which have a tetragonal internal structure show very interesting refractive properties,

which are well worth considering. These interesting effects occur, moreover, with the great number of crystals which are included in two other crystallographic systems—the hexagonal and the trigonal, while in the remaining systems,
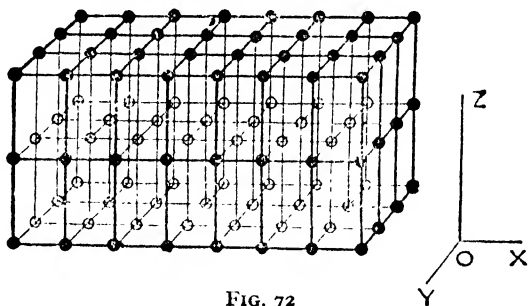
the orthorhombic, monoclinic, and anorthic, the behavioui to light is even more complex.

Returning to the typical crystal structure diagram of a tetragonal crystal (Fig. 72), it will be observed that light which is travelling in the direction OY will be exerting its electric forces in some direction in the plane ZOX, and these forces will have, in general, components in the two directions, OZ and OX. Now in the OX direction, the electrons are twice as dense as they are in the OZ direction, and so we shall expect the crystal to treat the two component vibrations (OX, OZ) in an appreciably different manner. This does, in fact, occur, and, in general, the two components are transmitted with different velocities and have different refractive indices. This difference in the treatment of the two component vibrations will also extend to any other properties of the crystal which are dependent on the electron arrangements. Thus, for example, some crystals are found in which the one vibration is absorbed far more rapidly than is the other. Tourmaline is the commonest example, though it is trigonal.

A further insight into the phenomenon is obtained by considering the wave surface of light emitted by a luminous point

embedded in a portion of the crystal. We will suppose the luminous point to emit suddenly a wave of light equally in all directions, and then we will consider the position of the wave a short time after, when it has spread out into the crystal. Imagine, for example, that we take an instantaneous photograph of the position of the front of the wave (Fig. 73).
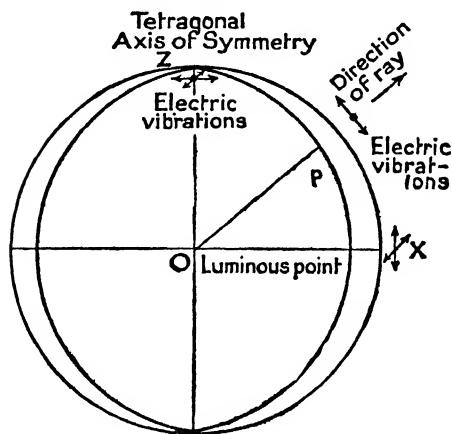
FIG. 73

A ray going from the luminous point in the direction OZ will have its electric forces acting in the direction shown, i.e. in the horizontal direction parallel to OX and 'in-and-out' of the paper. Since the direction OZ is the direction of the tetragonal axis of symmetry, the speeds and therefore the distances reached in equal times by these two components will be equal.

A ray going from the luminous point in the direction OX will have its electric forces having components which are, respectively, vertical and 'in-and-out' of the paper. Since, in the latter component, the electric forces are in the same direction as in the previous example, the distance this component will have gone from O will be the same as that of the two components belonging to the ray which is going along OZ. Moreover, a ray going in *any* other direction OP will also have one component of its electric force which is 'in-and-out' of the paper, so that this component will have reached the same distance from O as the similar component of the ray OX. Thus the wave part of this component will be a sphere surrounding the luminous point.

The second component of the ray OX is vertical and is, therefore, in the direction of different electron concentration. Its speed will, therefore, be different from that of the other component and may be greater or less according to the characteristics of the crystal. The second component for the ray OP will be at an angle with the vertical as shown, and therefore it also will be propagated with a speed which is different from that of the 'in-and-out' component; but since it is at an angle intermediate between the vertical component of OX and the horizontal one of OZ, its speed of propagation, and therefore its distance from the origin, will be intermediate between the two. In actual practice the wave front of this component is an ellipsoid touching the sphere at two points on the line OZ, which is called the optic axis of the crystal. This ellipsoid is sometimes larger and sometimes smaller in the directions OX and OY than the sphere.

The component whose electric force is perpendicular to the tetragonal (trigonal or hexagonal) axis is called the *ordinary* ray because its refractive index is independent of the direction of the ray. Its plane of polarization is, as has been agreed on page 117, at right angles to the direction of its electric vibration, so that its plane of polarization is the plane containing the tetragonal axis—i.e. the optic axis—and the ray.

The component whose electric force is partly parallel to the tetragonal axis is called the *extraordinary* ray, because its refractive index varies with the angle between the ray and the optic axis. The largest difference in refractive index between the ordinary and the extraordinary ray occurs when the rays are moving perpendicularly to the optic axis. The plane of polarization of the extraordinary ray is at right angles to the direction of its electric vibration and so is at right angles to that of the normal ray.

§ 64. Calcite

The most important example of doubly refractive substances is that of the trigonal crystal calcite, whose

*E

ordinary ray has a constant refractive index of 1·66, while that of the extraordinary ray varies from this value to 1·48 when the ray is travelling perpendicularly to the optical axis.

Natural crystals of calcite occur in a very large variety of forms, which readily cleave in three directions, forming rhombs as in Fig. 74. The exact shape of these cleavage rhombs is not of great importance for our present theory, but it is worth noting that when the reader finds one he will discover that all the faces of the rhombs will be parallelograms with
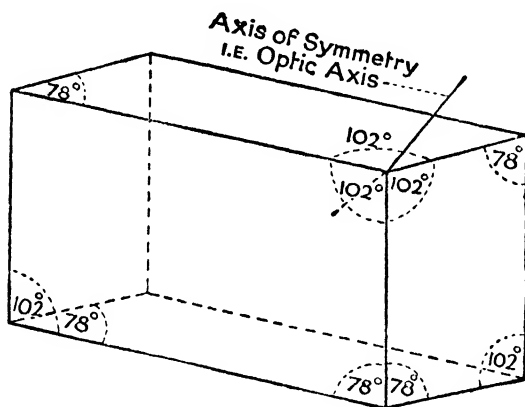
FIG. 74

angles of 78° and 102° as shown. At two of the corners of the rhomb all three face angles will be found to be 102°, and the direction of the optic axis of the crystal is that of the line, which makes equal angles with the three faces, through either of these corners.

We call this particular direction the direction of the optic axis of the crystal, because it is the direction of the trigonal axis of symmetrical arrangement of the atom units of the crystal itself. It is not, for example, the corresponding line drawn through one of the corners where three angles of 78° meet, because, although this does appear in the simple rhomb to

be the axis of trigonal symmetry, this is not so when the internal molecular arrangement is considered.   The optic axis, or the trigonal axis of symmetry, must not be confused with the line joining the two blunt corners of the rhomb, as the direction of this line clearly depends on the chance size of the crystal specimen.

Fig. 74 shows that rays which pass normally through the faces of the rhomb will be at a very considerable angle with the optic axis and so will have their vibrating electric forces split up into two components travelling at different speeds.

Ordinary
Extraordinary

t
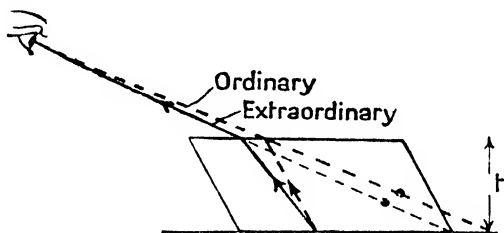
FIG. 75

These two components will, therefore, have different refractive indices, and a ray from a luminous object at the base of the rhomb will emerge into the air at two different angles as shown in Fig. 75.   This diagram should also make clear that the eye will see two images, due respectively to the ordinary and the extraordinary ray, which will naturally be situated at depths of $\dfrac{t}{\mu_0}$ and $\dfrac{t}{\mu_e}$.

The rays diverging from these two images are plane polarized at right angles to each other.   They show very interesting effects if they are made to pass through another calcite block similar to the first, but which can be rotated about the normal to the face.   In general each of the plane polarized beams will give two images whose intensities vary, as the upper block is rotated, in the manner we should expect if we resolved the initial plane polarized beam up into two components at right angles.

§ 65. The Polarizing Nicol Prism

This is the most important device we encounter in the study of polarized light. It is made by first cutting a natural rhomb of Iceland spar with a very thin diamond or carborundum wheel, then polishing the surfaces and reuniting with a film of Canada balsam. If the slice is cut at exactly the correct angle with the optic axis, and if the rhomb is of the correct dimensions, the ordinary ray with its large refractive index ($1 \cdot 66$) will be totally reflected from the Canada balsam layer ($\mu = 1 \cdot 54$) and will ultimately be absorbed by the black inside walls of the case in which the rhomb is mounted. On account of the lower refractive index of the

Balsam layer
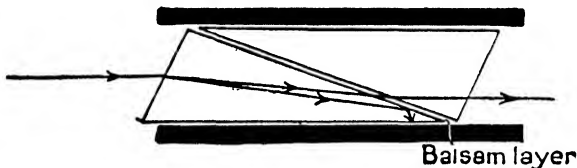
FIG. 76. THE NICOL

extraordinary ray, $\mu \sim 1 \cdot 48$, this will be directly transmitted through the prism and will emerge on the further side parallel to its original direction (Fig. 76).

As a practical tool the Nicol is invaluable, since it can be used to obtain plane polarized light in large intensity completely free from the other component of the original unpolarized light. The Nicol is equally important in the detection of plane polarized light, because if it is rotated round the direction of the ray a position will be reached at which the incident ray will have no component in the direction transmitted by the Nicol, so that complete blackness will result.

§ 66. The Analysis of Polarized Light

We have already remarked that if we are considering only light of one colour the magnitudes of the electric forces involved

are given by two equations having the same values of p — the angular frequency. We have also remarked that with unpolarized light there is an irregular phase difference between the two components in any two phases into which the electric forces can be resolved, whereas with polarized light there is a constant phase difference involved. The most general equations we can use to represent the electric forces in polarized light are therefore:

$$x = x_0 \cos pt,$$
$$y = y_0 \cos (pt + \delta).$$

The locus of the position an electron would occupy if it followed these forces will therefore be found, in suitable units, by eliminating the variable t from these two equations. This has already been done and the following result obtained:

$$\left(\frac{y}{y_0}\right)^2 - 2\frac{xy}{x_0y_0}\cos\delta + \left(\frac{x}{x_0}\right)^2 = \sin^2\delta$$

(see equation, page 113).

This represents in general an elliptic vibration on skew axes, which degenerates to an elliptic vibration on normal axes if $\cos\delta = 0$, to a circular vibration if $x_0 = y_0$ and if $\cos\delta = 0$, and to a straight line vibration if $\cos\delta = 1$.

Suppose we are now faced with the problem of the analysis of polarized light; we might, for example, be supplied with light from a hole in the wall and asked to determine if the light is or is not completely polarized, whether it is elliptically, circularly, or plane polarized, what are the angles between its axes and the horizontal, and what are the ratios of the two axes of the ellipse. The natural method of dealing with such a problem is to consider first the possibility of its being plane polarized light, then circularly polarized, and finally elliptically polarized. This natural method means, however, three separate investigations, or three separate discussions, and so it is best to consider only the one possibility that the light is elliptically polarized, since this one treatment includes in itself both plane polarized and circularly polarized light. We will, therefore, regard the light coming through the hole

in the wall as being composed of two components—the one with its electric force vertical and the other horizontal. The corresponding equations are then:

$$x = x_0 \cos pt \text{ and } y = y_0 \cos (pt + \delta),$$

and the locus of the corresponding electron motion will be an ellipse with its axes to the horizontal at the angle θ as in Fig. 77.

This same ellipse could have been made up by forces acting along the two lines OM and ON if these forces were given by the equations:

$$n = n_0 \cos pt \text{ and } m = m_0 \cos \left(pt + \frac{\pi}{2}\right),$$

providing m_0 and n_0 had been chosen correctly. This can be checked by eliminating (pt) between the two equations, and thus showing that it really does represent an ellipse whose axial lengths along OM and ON respectively are m_0 and n_0.
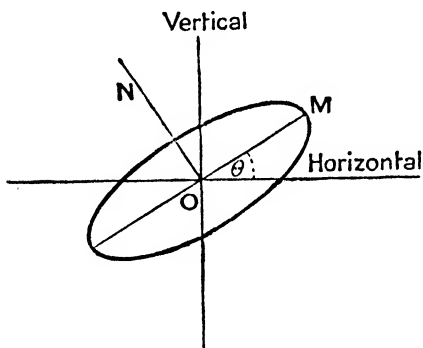
FIG. 77

In our experiment we will now pass this light through a plate of crystal cut so that the rays are passing perpendicularly to the optic axis. The one ray will go somewhat faster than the other, *and if the plate is exactly the correct thickness* the one will gain in phase by an amount c and the other by $\left(c + \frac{\pi}{2}\right)$, so that the phase difference will be exactly annulled.

The equations when the rays emerge from the crystal plate will then be:

$$n = n_0 \cos \left(pt + c + \frac{\pi}{2}\right) \text{ and } m = m_0 \cos \left(pt + \frac{\pi}{2} + c\right)$$

The locus of this is clearly the straight line $\frac{n}{n_0} = \frac{m}{m_0}$ making

an angle with OM whose tangent is $\frac{n_0}{m_0}$, which means that the electrons in the path of the ray will vibrate in a straight line and that the light can be stopped completely by a single Nicol

if it is set at this angle $\tan^{-1}\frac{n_0}{m_0}$ with the direction of vibrations in the crystal slice. In an actual experiment the crystal plate should have recorded on it the directions of vibrations—or of polarization—of the two components, and the Nicol should also have recorded on it the direction of the vibrations which it stops completely. The ratio of the two axes of the ellipse is thus simply the tangent of the angle between these two markings.

The directions of vibration which are recorded on the crystal plate will also represent the directions of the axes of the original elliptical light, so that we have determined the angles between the axes of the polarized light and the horizontal, and the ratio of the two major axes of the ellipse.

Finally, if the light were circularly polarized $\frac{m_0}{n_0}$ would be unity, and the angle of the axes of the ellipse indeterminate, while if $\frac{m_0}{n_0}$ were zero or infinity it would merely mean that the vibration was plane polarized.

The experiment would be conducted by passing the beam of light first through the crystal plate and then through the analysing Nicol. The crystal plate would be rotated round till the field of view was at its darkest, and then the Nicol rotated and set in its darkest position. This process would be repeated until both the crystal plate and the Nicol were in their darkest positions. The transmitted light should then be exactly zero unless it contained some unpolarized light, and, if it did, an estimate could be made of the relative proportions, by focusing the resultant light into a suitable photometer or photo-electric cell.

One difficulty regarding the analysis of the light is encountered in connection with the production of the crystal plate of the correct thickness to make a 90° phase difference, because the necessary thickness of the plate varies considerably with the wave length of the light involved. Babinet devised an arrangement of two quartz wedges, so cut that the vibration which gains on the other in the first wedge loses in traversing the second wedge.
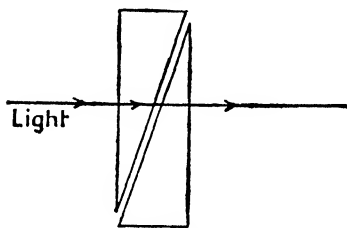
FIG. 78

From Fig. 78 it will be seen that a ray traversing the middle of the wedge, where the two thicknesses are equal, will have zero phase difference produced between its components. At increasing distances from the centre, the phase differences will vary progressively, so that it should be possible to adjust a tiny slit opposite that portion where the phase differences are suitable, and to employ this slit and quartz wedges to replace the quarter wave plate that has, so far, been considered.

Some arrangements of this Babinet compensator are of the type described, but others omit the single slit which we have suggested, and replace it by a system of rulings on the wedge. The eye is then concentrated essentially on that part of the wedge, which corresponds to 90° phase difference being introduced.

CHAPTER VI

PHOTOMETRY

§ 67. Introduction and Standard Sources. § 68. Definitions. § 69. The Relation between Photometric Quantities. § 70. Visual Effects. § 71. Optimum Room Illumination. § 72. Survey of Photometric Methods. § 73. Luminous Efficiency of Different Colours. § 74. Photoelectric Cells. § 75. Photoelectric Cells of the Copper/Copper-Oxide Type.

§ 67. Introduction and Standard Sources

Light is a form of electromagnetic radiation, and involves a transfer of energy in the direction of propagation. It is, therefore, natural to attempt to describe the intensity of light by the energy in ergs per second which crosses unit area perpendicular to the ray. This procedure is not generally adopted, for two essential reasons.

Firstly, the actual amount of light energy with which we are usually concerned is exceedingly small, since the eye is a far more sensitive detecting instrument than any other apparatus ordinarily used for the purpose. For example, the limit of detection by the eye is about $1 \cdot 7 \times 10^{-16}$ watts (at the wave length of maximum visibility) while the radiometers used for measuring the heat emission of stars have a detection limit of about 1×10^{-10} watts.

Secondly, the visual effect produced in the eye by a constant amount of light energy varies very markedly as the wave length, frequency, or colour of the light is altered; this will be discussed later.

In view of these difficulties attending the measurement of light quantities in absolute units, an arbitrary system of units was proposed, and in 1860 the British candle was adopted

as the legal unit. Now the most important qualifications of a standard are that it should be both constant and reproducible, and the British legal candle was found to be unsatisfactory in each respect. Vernon-Harcourt, therefore, devised another lamp, and showed that it was accurately constant in its behaviour. This lamp burns pentane (C_5H_{12}) and is called the Vernon-Harcourt pentane lamp; it is not easily made in small sizes and was, therefore, constructed to produce about ten times the light of a standard candle. The modern standard candle is now defined legally as being one-tenth of the candle power of the pentane lamp, so that all light sources are now referred ultimately to the pentane lamp, and not to the old spermaceti candle. In other countries other lamps were used as standards until 1909 when they were all expressed in terms of the single international unit of one-tenth of the candle power of the Vernon-Harcourt pentane lamp.

Attempts are now being made at the different National Physical laboratories of the world to obtain a greater accuracy in the specification of the standard light source, and to replace the Vernon-Harcourt lamp by a perfect black body at the freezing-point of platinum. The candle-powers determined experimentally from unit area of such a source are at present 59·1 candles $\pm$ 0·5% in England, 58·86 in U.S.A., and 58·78 in Strasburg. This accuracy is just not sufficient to justify any legal alteration in the standard by international agreement.

Another recent development has been the use of specially designed electric lamps as sub-standards, both in industry and even in the National Physical Laboratory itself. These sub-standards have been shown to be very constant indeed for a considerable fraction of their total life, although they are very sensitive to variations of voltage (1% change of voltage producing 3·6% change in candle-power). On account of the convenience of these sub-standards, the Vernon-Harcourt pentane lamp is practically never used now either in industry or in the laboratory, and even in the National Physical Laboratory itself it is only used occasionally to check the sub-standards.

§ 68. Definitions

(1) *The International Candle.* This is defined by the statement that 10 international standard candles would be identical with the Vernon-Harcourt pentane lamp, measured in the standard direction and in the standard manner. The international standard candle is, therefore, an abstraction, and it is usual to consider it theoretically as being a source which is constant in all directions.

(2) *The Lumen.* This is the unit of luminous flux, and is the quantity of light emitted per second in each unit solid angle surrounding a uniform point source of one international candle-power. (Symbol L.)

(3) *The Illumination or Intensity of Light on a Surface.* Symbol I. This is the quantity of light falling per second on unit area of the surface. It may be measured in terms of several alternative units, as follows:

> The Lux. This is one lumen of (light energy per second)* falling uniformly on one square metre.
> The Phot. This is one lumen of (light energy per second) falling uniformly on one square centimetre.
> The Foot-candle. This is one lumen of (light energy per second) falling uniformly on one square foot.
> We can easily deduce from these definitions that one lux is the intensity of illumination produced on a perpendicular surface at a distance of one metre from a standard candle; one phot the intensity produced by a standard candle at one centimetre; and one foot-candle that produced at a distance of one foot.

(4) *Brightness.* The brightness of a source, or of an illuminated area, is proportional, in the direction considered, to the

---

* Notice that a lumen is a unit of light energy per second, and is not simply a unit of light quantity. The number of lumens falling on a given surface is therefore a measure of the rate of transfer of luminous energy, in much the same way that an ampere is a unit of electric current $\frac{dq}{dt}$, and is not the unit of quantity of electricity—which is the coulomb.

effective candle-power per unit area of the surface. The unit of brightness is called the lambert, and is the brightness of a perfect diffusing surface radiating one lumen per square centimetre.

If light from a standard candle at a distance of one metre falls normally on to a perfect white matt diffusing surface, each square metre of the surface receives—and therefore radiates —one lumen of light energy per second. The brightness of the surface is, therefore, 10^{-4} lamberts, or 0.1 milli-lamberts.

It is advisable to draw a clear distinction between 'illumination' and 'brightness.' A piece of white paper and a piece of black velvet lying side by side may be equally illuminated, but their brightnesses will be very different. The brightness of a surface is proportional to the product of the intensity of illumination and the diffusion coefficients of the surface.

§ 69. The Relation between Photometric Quantities

(a) In view of the fact that the intensity of light has been defined as the quantity of light falling per second on unit area, it should be related to the candle-power, C, of the light source, the distance, D, between the source and the screen, and the angle θ between the ray and the surface normal by the equation:

$$I = \frac{C}{d^2} \cos \theta.*$$

---

* Quantity of light crossing A′ (Fig. 79) per second is the quantity crossing A, and this should be equal to the total number of lumens $4\pi C$ emitted by the lamp times the ratio of the area A to the total area of the sphere $4\pi d^2$, i.e. the quantity of light crossing A′ is
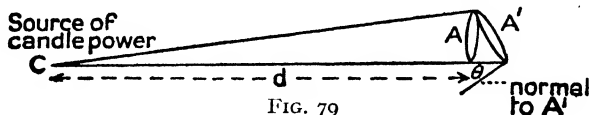

FIG. 79

$$4\pi C \frac{A}{4\pi d^2}$$

$$= C\frac{A}{d^2}$$

$$= C\frac{A'}{d^2} \cos \theta.$$

(*b*) In all the ordinary instances in which a surface is illuminated by several independent sources of light, the intensities of the light add up directly, and

$$I = I_1 + I_2 + I_3 + \text{etc.}$$

The only exception to this rule occurs when there is a *definite phase relationship* between the sources of the type considered in connection with the 'interference of light.' The rule is correct even if the separate sources are of different colour, even though the resultant *colour* of the mixture may appear to us to be markedly different from that of any of the original components.

(*c*) The apparently complicated method of defining our units of illumination as the number of lumens falling on unit area is due essentially to the fact that ordinary sources are usually very far from being symmetrical in all directions. Thus a practical lamp will behave as if it were of larger candle-power in one direction than in another. Because of this difficulty, the modern tendency is to express all light quantities, equations, and definitions in terms of the unit of light energy crossing an area per second—which is the lumen.

Thus the mean spherical candle-power of a practical lamp is the total light energy it emits per second divided by the corresponding quantity which would be emitted by a standard candle if that were uniform in all directions. The mean spherical candle is thus:

$$C = \frac{\text{total lumens emitted}}{4\pi},$$

since 4π lumens come from one uniform standard candle.

In a similar way, the candle-power of a lamp measured in any particular direction is the (number of lumens included inside a solid angle $\delta\omega$ in the required direction) divided by the corresponding number ($\delta\omega \times 1$)' emitted by a unit standard candle. Thus:

$$\text{Candle-power in any direction} = \frac{dL}{d\omega}.$$

§ 70. Visual Effects

It is very important to notice that the quantity I—the intensity of the light—does not represent directly the *visual effect* produced in the eye. That is, the magnitude of the resultant effect produced in the mind is not simply proportional to the intensity of the light or to the total quantity of light energy received per second.

Verifications of the formula $I = \dfrac{C \cos \theta}{d^2}$ cannot therefore be made by estimating visually the intensity of light, but must be performed with instruments recording the actual energy per second of the light beam or by comparison experiments, using, for example, the Bunsen grease spot photometer. In the first type of experiment, most instruments must be calibrated, so that in most experimental work we effectively test the validity of the inverse square law by relying on our conviction that the intensity of light on a screen is proportional, if the distances are kept constant, to the number of separate standard candles which could replace the lamp considered.

It is exceedingly difficult to obtain any estimates of the magnitude of the visual effect produced by light—our sensation of brightness is very dependent on our own immediate past history, and a normally dull object can appear quite bright if we view it immediately after coming from a dark room. The only definite information that is available concerns the magnitude of the change in intensity dI on a portion of a uniform field of view, which can just be detected by the eye. If we call this small change in visual effect dV the experimental result is that

$$dV = \frac{1}{A} \frac{dI}{I},$$

where A is approximately constant and is called the 'Fechner fraction.'

Although it does not seem very justifiable to determine the

total visual effect of any luminous object by integrating the quantity dV, it is frequently stated that

$$V_1 - V_2 = \frac{1}{A} (\log I_1 - \log I_2).$$

The general probability of the approximate truth of this law can, however, be seen by considering the visual effects produced under different illuminations by two adjacent parts of an object (e.g. the face and hair of a person seen in poor light and again in strong sunlight). The intensity of light re-emitted by the two parts of the object when exposed to the general illumination I′ can be written as f_1I′ and f_2I′, where f_1 and f_2 represent the fraction reflected or diffused by the two parts of the object considered.

Thus the difference between the two visual reactions will be

$$= \frac{1}{A} \log (f_1 I') - \frac{1}{A} \log (f_2 I')$$

$$= \frac{1}{A} \log \left(\frac{f_1}{f_2}\right),$$

The corresponding difference between the two visual effects when the general illumination is I″ will also be

$$\frac{1}{A} \log \left(\frac{f_1}{f_2}\right),$$

so that, if the law is correct, we should observe the same contrast between different parts of objects whatever is the general illumination.

The importance of this may be judged from the fact that if the visual effect had been proportional to the intensities of the light received instead of to the logarithm of the light intensities, the contrast between the two parts of the object would have been $(f_1 - f_2)$I′. Thus if I′ had been 5,000 foot-candles—as it is for daylight, and if I″ had been 1 foot-candle —for artificial light, the contrasts would have been $(f_1 - f_2) \times 5,000$ and $(f_1 - f_2) \times 1$ respectively, so that the object could have looked entirely different under the two conditions.

We have already mentioned that the 'Fechner fraction' A

in the equation for the smallest perceptible change in visual effect $d\mathrm{V}$,

$$d\mathrm{V} = \frac{\mathrm{I}}{\mathrm{A}}\frac{d\mathrm{I}}{\mathrm{I}},$$

is only constant to the first order of accuracy. At very low intensities we need a comparatively large *percentage* change (5–10%) in the illumination of a part of a field of view in order to make a perceptible change in the visual effect. Similarly our eyes are not very sensitive to changes of intensity when the illumination is very large indeed. Between these two extremes the minimum value of A occurs at which we can detect a sharp change of light intensity as small as 1%. In view of the existence of this minimum value of A, and therefore of the corresponding maximum value of the sensitivity of the eye, it is advisable to conduct photometric measurements with light intensities between 1 and 1,000 foot-candles.*

The results obtained by König and Brodhun for the smallest detectable percentage change in the brightness of part of a field of view are as follows:

TABLE V

Field Brightness in 10^{-4} lamberts	40,000	4,000	400	40	4	0·4	0·04
% Change detectable	3·6%	1·9%	1·76%	1·78%	2·98%	4·8%	12·3%

(A perfect white screen illuminated with 1 ft.-candle or 10·8 metre-candles has a brightness of 10·8 × 10^{-4} lamberts.)

Other workers claim to be able to detect a change of only 1% and even 0·7% in certain cases when working under the optimum possible conditions. These differences are probably explained partly by the character of the line of demarcation

* For more details refer to the *Dictionary of Applied Physics*, or to the Report of the Medical Research Council on the 'Physiology of Vision,' by R. J. Lythgoe (His Majesty's Stationery Office, 1926).

between the two parts of the field and partly by the area of the field itself, since the smallest percentage difference in brightness that can be detected between the two halves of a circular field of view is inversely proportional to the diameter of the field if this is smaller than about ½ cm. in diameter, and is viewed from the nearest distance of distinct vision.

§ 71. Optimum Room Illumination

The optimum illumination for rooms is not simply that illumination at which the eye is most sensitive to small changes in brightness, for it is certainly dependent on many additional factors, such as the character of the work considered; the glare and the contrasts; and the extent to which the eye is allowed to remain adapted to the one illumination.

There is, however, a definite direct connection between the output and accuracy of fine work (such as setting type by hand) and the illumination. Employers should consider this if they wish to combine a maximum of output at a minimum of cost with a high degree of comfort to their employees. Thus the work performed at illuminations of 1 and 7 foot-candles is only 76% and 88% respectively of the maximum achieved with daylight or with 25 foot-candles. In addition, the errors under the poor illumination are more than twice as many.

Clearly the optimum illumination for drawing on a white paper will be less than for sewing black thread on a black cloth, so that we are concerned not so much with the actual room illumination as with the brightness of the object (or of the image which is formed on the retina). Thus a uniform grey sheet which reflects only 10% of the incident light requires ten times the illumination of a white sheet to appear of the same brightness. If the character of the work to be performed is the same, the illumination required is inversely proportional to the reflection ratio.

It is by no means easy to calculate the most profitable

illumination to provide in industry, but the values usually recommended are: 30–60 metre-candles (approximately 3–6 foot-candles) for reading or writing; 10–20 metre-candles (approximately 1–2 foot-candles) in foundries where no fine work is done; and 80 metre-candles (8 foot-candles) in drawing offices where very great precision is needed and where expensive labour is employed.

An increase in illumination can affect the eye in three ways. It can produce an alteration in visual acuity, so that the observer can read books or print for which the individual letters subtend smaller angles at the eye. It can affect the speed of recognition of the letter, or of the recognition of a small gap in Landolt's broken white ring, and it can delay the natural development of eye fatigue.

With regard to visual acuity it has been shown that at an illumination of 0·016 metre-candles (0·0016 foot-candle) the eye can perceive letters 2mm. in size at one metre. The acuity first increases rapidly with brightness and then much more slowly, so that the visibility limits (at one metre) are letters of 1·6 mm., 1·3 mm., and 1·0 mm. with intensities of 1·5, 16·7, and 5,400 metre-candles respectively.

In another investigation of the same type the following data were obtained:

TABLE VI

Change in Intensity in foot-candles	0·001 to 0·1	0·1 to 1·0	1·0 to 5·0	5·0 to 20·0
% Increase in Acuity	490%	67·7%	43·6%	8·2%

A further investigation showed that a higher illumination is required if the subject exhibits only small contrasts, so that, although little improvement in visual acuity for black print on white paper occurs with illuminations of more than 4 foot-candles, there is still a considerable gain with dull print on a brown background up to above 25 foot-candles.

The speed of reading has been shown to increase very

rapidly up to about 0·5 foot-candles, and is steady above 10 foot-candles with subjects showing small contrasts.

The disadvantage of poor systems of lighting is not observed so readily with a fresh eye, but becomes noticeable because of the inability of the eye to maintain its efficiency. Poor illumination, therefore, means greater eye fatigue, and a more rapid diminution of efficiency.

The diminished efficiency of a fatigued eye is exemplified by the fact that patients looking at test cards get their best vision at the beginning, and that the longer they look at the cards the more mistakes they make.

§ 72. Survey of Photometric Methods

Up to the present we have referred to the candle-power of lamps and to intensities of illumination without discussing the experimental methods by which they are determined.
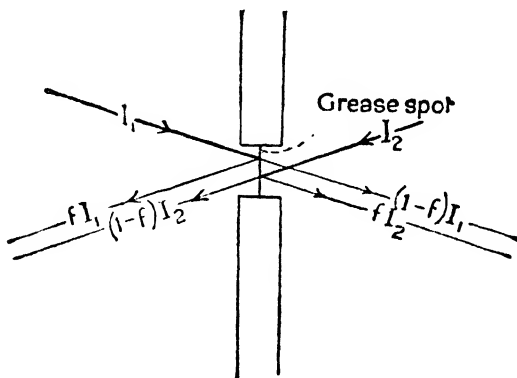

FIG. 80

A very large variety of instruments have been devised for these purposes, and can be illustrated by the Rumford shadow photometer, the Bunsen's grease spot photometer, the Lummer-Brodhun photometer, the flicker photometer, and by the modern photo cells of the normal and of the copper-copper

oxide type (which are discussed on page 147). Since most of these instruments are discussed completely in the usual text-books it is only necessary to mention a few isolated points.

(a) *The Rumford Shadow Photometer.* This has very poor accuracy, since—owing to the finite penumbra—there is no sharp line of demarcation between the two shadows. It is unsatisfactory for colour.

(b) *The Bunsen Grease Spot Photometer.* The theory of this is not usually completely understood. If I_1 and I_2 are the intensities of illumination on the two sides respectively, and if a fraction f is diffused backwards from the grease spot, then the amount penetrating the spot will be $(I_1 - fI_1)$ and $(I_2 - fI_2)$ respectively, if no light is absorbed by the grease spot (Fig. 80).

Thus, on the one side, the intensity on the grease spot will be
$$(I_1 - fI_1) + fI_2.$$
We shall then be unable to distinguish this from the matt background, which diffuses I_2 if
$$(I_1 - fI_1) + fI_2 = I_2.$$
The normal treatment then goes on to express this equation as
$$I_1(1 - f) = I_2(1 - f),$$
and so proves that if the grease spot is indistinguishable from the background, then I_1 must equal I_2.

The important point to consider is, however, *how accurately does* $I_1 = I_2$?

When we were discussing the 'Fechner fraction' it was stated that, under good conditions, and with intensities of about 1 foot-candle, normal observers would observe equality in two parts of a field of view over a range of about 1%. It therefore follows that
$$(I_1 - fI_1) + fI_2 = I_2 \text{ with a probable error of } \tfrac{1}{100} I_2.$$
So $\quad I_1(1 - f) = I_2(1 - f) \pm \tfrac{1}{100} I_2$
$$I_1 = I_2 \pm \tfrac{1}{100} \frac{I_2}{(1 - f)}.$$

Thus we can only hope to determine if $I_1 = I_2$ to an accuracy

of 1%, and, *if (f) is at all large, as it will be for a poor grease spot, the accuracy will be much less than* 1%.

A considerable improvement in accuracy can be made by taking a large number of *independent* readings (accuracy of mean $\dfrac{1\%}{\sqrt{n}}$), or by noticing the distance of the variable lamp when the grease spot is respectively just perceptibly brighter and just perceptibly darker than the background.

With practicable grease spots an appreciable fraction of light is absorbed, and, in consequence, if there is perfect equality between spot and background on the one side, then on the other the spot will be materially the darker.   The adjustment should, therefore, really be performed so that *the contrast between spot and background is the same on both sides.*

(c) *The Lummer-Brodhun Photometer.* This instrument shows three great advantages over Bunsen's grease spot.   It has very sharp lines of demarcation between the different parts of the two fields.   It has a value of *f* which is appreciably near to zero, and it is completely free from the difficulty of absorbed light.

Like the grease spot, the Lummer-Brodhun photometer is unsatisfactory for sources of different colour.   It is generally employed in the National Physical Laboratory for comparing the candle-power of lamps as it gives greater accuracy than any other type of photometer so far devised.

(d) *The Flicker Photometer.* This instrument is probably as important as, or even more important than, the Lummer-Brodhun photometer, because it is capable of almost the same accuracy, and can, in addition, be used for comparing the brightness of lights of different colour.   A number of different modifications of the instrument have been employed, but they all use the same basic idea of presenting to the eye in rapid alternation two surfaces illuminated respectively by the two lamps investigated.   If the speed of the alternations between these two is increased a sense of flicker will be felt at speeds of about 15–20 a second (the exact value increasing with the intensity of the light).   This flicker will disappear again at speeds of about 30 a second and a uniform field will

be seen. At frequency alternations intermediate between these two, it has been found that the sense of flicker disappears if the two intensities of illuminations are equal to within about 1%.
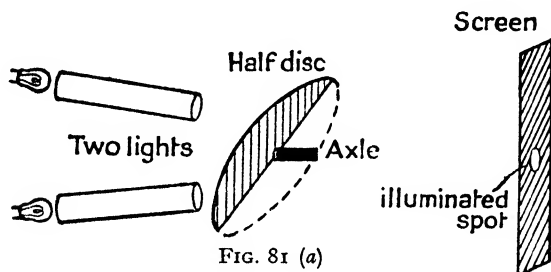
FIG. 81 (a)

The simplest experimental arrangement (outlined in Fig. 81 (a)) uses one screen illuminated by the two lamps, the light from which is cut off alternately by a suitable disk mounted on the shaft of an electric motor.
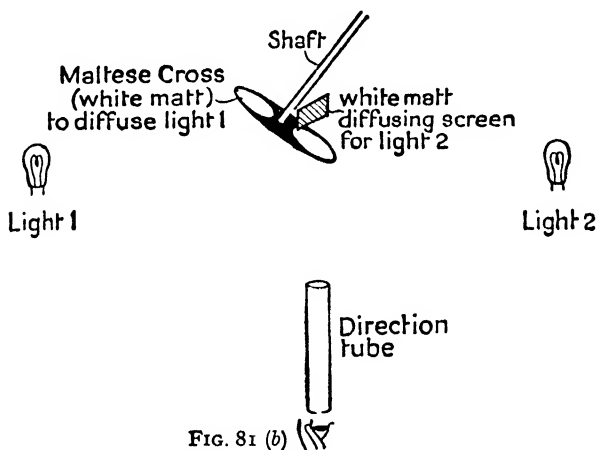
FIG. 81 (b)

Another arrangement is that portrayed in Fig. 81 (b). For details of this very important modification refer to the usual text-books.

§ 73. Luminous Efficiency of Different Colours

As has already been mentioned, the great importance of the flicker photometer lies in its capacity to compare the brightness of lights which differ in colour. It is found that, at slightly higher speeds than are required for light of the same
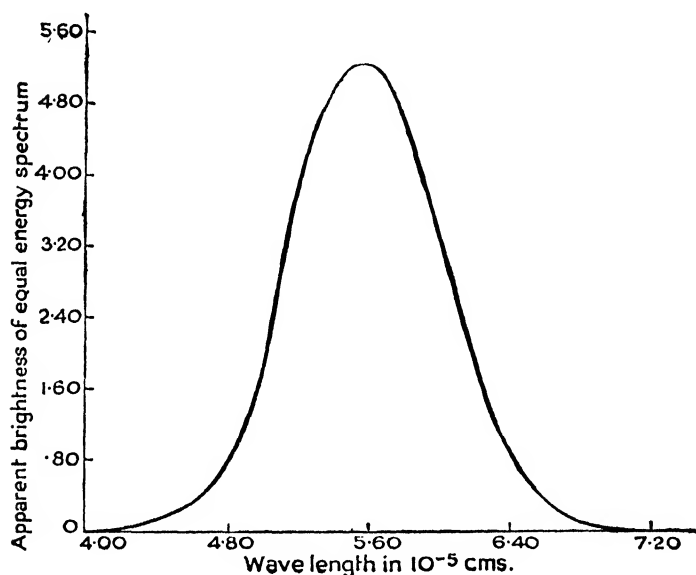
FIG. 82

The data for this curve are given in Table VIII, page 157.

colour (speeds between 25 and 30 alternations per second are suitable), the two colours disappear and give rise to a uniform colour with a sense of flicker superimposed, which can be made to disappear by adjusting the light intensities.   When this occurs the lights are said to be of the same apparent brightness, even though the actual energies (in ergs per second per sq. cm.) of the two may be very different.   Graphs can then be constructed connecting the wave-length—or colour—of the light with the energy in ergs per second per sq. cm. needed to

produce a constant apparent brightness, or correcting the colour with the apparent brightness of lights of the same energy.

It is often believed that, by this graph, we are claiming to compare things which differ in character as well as in quantity. This belief is clearly incorrect because we are comparing one quality only (viz. brightness) of the different coloured lights. In a similar way we compare the length of a table with the length of a metre stick without claiming that the table and the stick are identical in all respects other than that of length.

It does seem, moreover, that a definite physical meaning must be attached to the graph, since it is constant and is reproducible for any individual observer, and is, moreover, the same for all observers—providing they are not colour-blind.

In order to confirm that the flicker does disappear when the lights have the same visual effect, the following, rather inaccurate, experiments have been performed.

(a) Guesses have been made of the apparent brightness of the different lights. Such guesses are very variable, but have been made more reliable by taking the average of a very large number of attempts.

(b) The visual acuity (i.e. the smallest size print that can be read) has been determined for the different colours.

(c) The minimum intensity required to see objects (i.e. the threshold of vision) has been measured.

The latter method is, however, somewhat unsatisfactory, in view of the existence of the Purkinje effect, viz., that at very low intensities colour is not recognized. Thus, if an observer in a dark room is regarding a spectrum which is being uniformly reduced in intensity, a stage will be reached at which the red has disappeared, the orange has become yellow, and the green bluish. With still lower intensity the process continues until the whole of the yellow-blue field becomes a uniform bluish-grey.

The effect is explained by assuming that the cones of the retina alone are sensitive to colour but that they require much more light to excite them than rods, which are, therefore, responsible for vision at low intensities.

§ 74. Photoelectric Cells

An entirely new principle of light measurement has been developed recently in the photoelectric light-sensitive cells. The latest modification of these is of great importance in investigations of light of different colours, because it has a sensitivity curve (current produced per unit energy received/wave-length) which approximates very closely indeed to that of the normal eye.

This property, combined with the general convenience of the instrument, makes it appear certain that in a few years all other types of light intensity meters will be largely out of date except for very special purposes. A general discussion of these photoelectric cells is, therefore, desirable.

Three main types are classified under the term 'photoelectric cell.' The first of these—the selenium cell—uses the principle that the resistance of selenium decreases very considerably when illuminated. For the purposes with which we are concerned it is sufficient to reg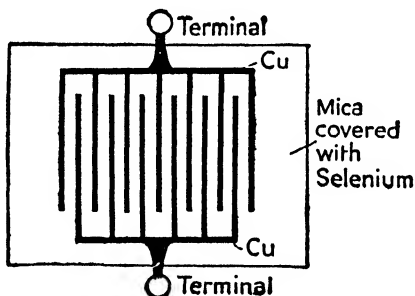ard this as a simple experimental fact without inquiring why selenium has this property. For efficient working the selenium layer should be so very thin that the light penetrates in appreciable intensity right through it, so that a practical arrangement might consist of a series of copper wires glued to an insulator as in Fig. 83, and the whole coated, in any convenient manner, with a thin film of selenium.

FIG. 83

Much more important instruments are the normal vacuum or gas-filled photoelectric cells.

These often resemble in external appearance ordinary wireless valves, and usually have a standard four-pin base. They

F

are generally represented diagramatically as in Fig. 84, because they contain two electrodes. The one consists of a thin film of potassium, caesium, or rubidium, which has been deposited over the larger part of the inside of the glass bulb, while the second electrode is small and is well insulated from the first.
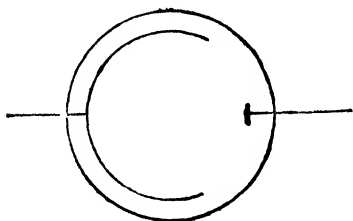

FIG. 84

Under the action of any light whose wave-length is shorter than a certain minimum (say red or orange) it is an experimental fact that electrons are ejected with slow velocities from the potassium or caesium surface. Once again this should be regarded here as a primary experimental fact, although it can be connected up with a large number of other interesting phenomena by the 'quantum theory.'

In the vacuum type cells these electrons cross the bulb without being stopped by any gas atoms and are attracted by the small electrode if this is maintained positive with respect to the larger electrode by an external battery. This battery, therefore, produces a current which is proportional to the number of these electrons produced per second, and thus proportional to the intensity of the light.

The gas-filled cell contains an inert gas, such as argon, at a very low pressure, such that an average electron produced by the caesium surface makes only a very few (4–8) collisions as it crosses the bulb towards the small electrode. If, however, the voltage between the two electrodes is high enough, an electron will gain so much energy between each collision that it can ionize the gas atom when it does make an encounter, and thus produce two more ions. These two ions will, in turn, produce still more, so that the current might be expected to rise to very large values with the production of a spark. This will not, however, occur if the pressure of the gas and the voltage between the electrodes is adjusted to special values,

although an alteration of this adjustment may cause a spark, or an arc, to occur and so ruin the cell. It is then found experimentally that at the correct gas pressures and at the correct voltages the current is accurately proportional to the number of primary electrons and therefore to the intensity of the light illumination, but is greater than the primary electron current by ratios of the order of 100 : 1, and will amount to the order of 10^{-6} amp. per lumen.

Many interesting applications of these cells have been devised in which the output current from the photocell is amplified by the use of valves and relays.

§ 75. Photoelectric Cells of the Copper/Copper-Oxide Type

The most important types of photoelectric cell for use in photometric measurements are those of the copper/copper-oxide type.

Discs of highly purified copper are heated in air or in oxygen to a high temperature in an electric furnace. The discs partially oxidize and produce a surface layer of black cupric oxide, and an inner layer of cuprous oxide on the basic copper. The black cupric oxide layer is usually removed by rubbing with emery paper or by solution in sodium cyanide.

These units are the basic elements of commercial 'all-metal rectifiers,' as they exhibit a very much greater electrical resistance between the copper and a lead grid placed on the cuprous oxide when the copper is positive than when it is negative. Thus, with the copper at $+5$ volts relative to the lead, the resistance may be 16,000 ohms, while with the copper at -5 volts, it may be as low as 25 ohms.

This peculiar rectifying action is not completely understood, though it is presumably bound up with the fact that the crystals of the cuprous oxide have all grown in a unidirectional manner from the copper base, so that there is a cleaner, more intimate contact between the two bodies than would be possible with any other mode of formation.

These discs have, in addition, been found to be sensitive to

light, and can be used as photoelectric cells. For this purpose a fine grid of lead wire can be pressed into contact with the cuprous oxide (back-walled type), or a very thin film of gold, or similar metal, deposited on the oxide layer (front-walled type, Fig. 85).

In these front-wall copper/copper-oxide photoelectric cells the gold film is so very thin that it transmits a considerable
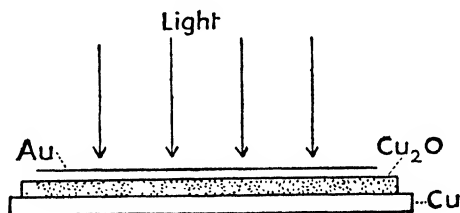

FIG. 85

part of the incident light to the cuprous oxide layer. We then have the further experimental result that when the light is absorbed by the cuprous oxide, electrons are ejected, just as in the ordinary caesium photoelectric cells. Here, however, the electrons are ejected only from the cuprous oxide into the gold, and not appreciably from the gold to the cuprous oxide; a small battery is therefore formed between the gold and the copper surface, which can pass through a connecting wire from the gold to the copper a number of electrons which is directly proportional to the intensity of the incident light.

The very great importance of the apparatus lies in the facts that no external battery is required and that the current, with intensities of 1–20 foot-candles and with areas of surface of about 5 sq. cm., is of the order of 10^{-5} of an amp. and so can be determined by robust, double-pivot, moving-coil micro-ammeters.

In addition, some photoelectric cells are available with a spectral sensitivity curve which is very close to that of the normal eye. This is shown in Fig. 86.

These copper/copper-oxide photoelectric cells are available separately, or permanently connected to moving-coil micro-

Cell output

Human eye response

Wave length x 10^{-5}cms.

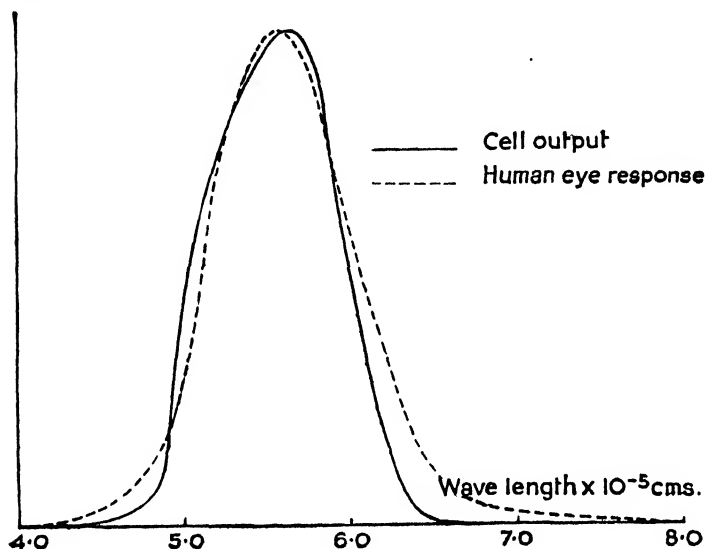4·0   5·0   6·0   7·0   8·0

FIG. 86

ammeters, and are calibrated to register directly in foot-candles, and can thus be used for comparing light intensities, and therefore candle-power of lamps, in a direct, simple manner quite independent of differences in colour.

CHAPTER VII

COLOUR

§ 76. Introduction. § 77. Three Types of Visual Receptors. § 78. Elementary Colour Specification. § 79. Interpretation of a Colour whose Blue, Green, and Red Brightness Factors are known. § 80. Hue and Saturation. § 81. Colour Mixing and Matching. § 82. Colour Mixing to produce White Light. § 83. Other Systems of Colour Specification. § 84. Warnings. § 85. References. § 86. Light Filters. § 87. Paint. § 88. Paint—Browns and Greys.

§ 76. Introduction

An attempt is now being made to lead the colour industries from their present state of individualistic chaos by providing an international scale and international units in which any new colour can be uniquely specified. This cannot, in any way, cramp the industry, or prevent the discovery or utilization of new shades. It does not forbid romantic new trade names being employed for either new or old shades; the international scale and units merely provide the means by which such new shades may be specified, so that any person who desires an accurate colour-match for any article can refer to its specification and write to a manufacturer for a similar colour without enclosing a sample and yet have perfect confidence that an *inaccurate* colour-match should *not* be delivered.

The subject of the colours of pigmented articles is very complex, as such colours depend to some extent on the brightness of contrasting objects as well as on the character of the illuminating light. We may start, however, by observing that all coloured objects are seen by light which they emit or diffuse, so that we may concentrate attention on the effects produced by radiation of different wave-lengths entering the

eye irrespective of the manner in which this light has been produced. We may therefore forget, for the present, any difficulties regarding colours produced by mixed paints or dyes.

Electromagnetic radiation can be detected as light by the eye only if its wave-lengths are between 4.0×10^{-5} cm. and 7.2×10^{-5} cm. (i.e. 4.0 and $7.2 \times 10^2 \mu\mu$ or 4.0 and 7.2×10^3 A.U.) Thus all coloured lights are formed by combinations in different intensities of radiations whose wave-lengths are between these two ranges.

§ 77. Three Types of Visual Receptors

We have already discussed (p. 143) the experimental determination of the intensities of monochromatic light of different
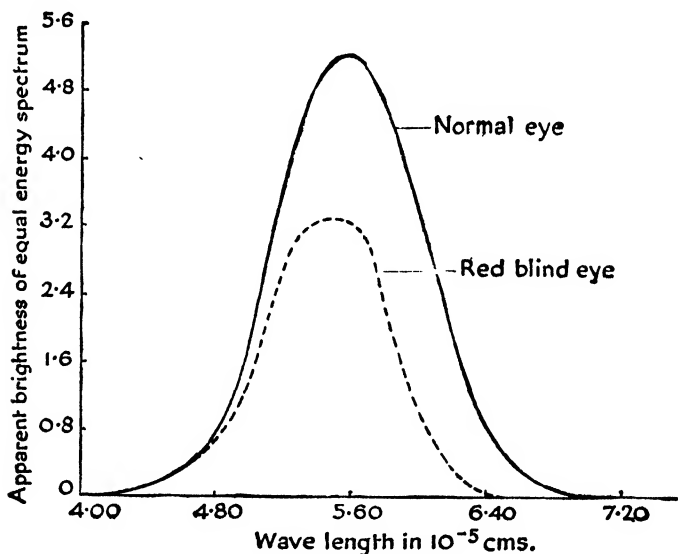

FIG. 87

The data for this curve are given in Table VIII, page 157.

wave-lengths which are needed to produce the same apparent brightness. These methods can be summarized as follows:

(1) Flicker photometer, (2) equal ease in reading, (3) apparent brightness, (4) minimum intensity to observe certain objects. From these experimental results a curve has been constructed showing the relation between the apparent brightness and the wave-length for lights of the same energy. This curve, which was shown in Fig. 82 (p. 143), is repeated in Fig. 87.

Now the phenomena of colour matching and colour mixing can be explained by the assumption that colour sensations result from the stimulation of three different types of receiving mechanisms in the eye. It is uncertain whether there are three types of cones in the retina, or whether each cone can be excited in three different ways.

The total impression of brightness registered in the brain when light falls on the retina is the sum of the contributions produced by each of these three types of receiving mechanisms. The curves of Fig. 88 show how the apparent brightness registered by each of these three receptors depends on the wave-length of the light, which is kept at constant energy. These curves show that, in general, one single monochromatic radiation will excite all three receptors simultaneously. (Naturally, the sum of the brightness-ordinates for each wave-length gives the normal visibility curve (Fig. 87) of the eye.)

It is convenient to assign names to these three receptors, and it is usual to call them the blue, green, and red receptors respectively. These names are perhaps a little unfortunate, since each receptor is stimulated to some extent by almost the whole range of the spectrum. On account of the small relative stimulation of the blue receptor, the dotted curve has been drawn with the brightness-ordinates increased by a factor of ten.

These curves are actually based, in a somewhat complicated manner, on the experimental results of König, Dieterici, and Abney on colour mixture, and they have been largely confirmed by experiments on colour-blind persons, although curves derived from such experiments are necessarily less accurate than are those based on colour mixing.

Slight differences exist between the sensitivity/wave-length curves of normal observers, which may be due to one or two of the three types of receiving mechanisms being different from, or less sensitive than, the normal.   This phenomenon
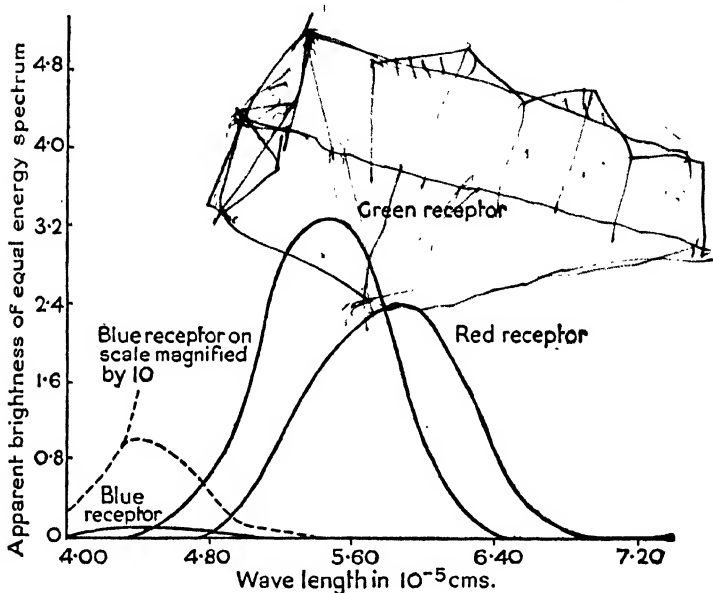

FIG. 88

The data for this curve are given in Table VIII, page 157.

occurs most commonly in the red or green, people suffering from it being called 'anomalous trichromats.' Particularly bad cases are known in which the person is completely blind in one or other of these receiving mechanisms.   The most common type of colour-blind person is called a 'protanope' and has the red visual receptor missing.   In 'deuteranopes' the green curve is missing, and in certain very rare cases—'tritanopes'—the violet curve is absent.   Colour-blindness is hereditary and is, according to its peculiar Mendelian law, much more common

*F

with men than with women. The above simple description of colour-blindness, as due to the absence or inactivity of one of the visual receptors, is known to be only a first approximation to the truth. It is more likely that in colour-blindness the spectral characteristic of one of the receptors is changed so that it becomes identical with one of the others.

We should naturally expect to obtain an *approximation* to the sensitivity curve of the red receptor by subtracting the sensitivity curve of a totally red-blind person from that of a normal eye. A comparison of the normal and the colour-blind (brightness/wave-length) curves of ·Fig. 87 shows that this is roughly true.

§ 78. Elementary Colour Specification

If monochromatic light falls on to the retina, all three receptors are stimulated to some extent and contribute towards the total brightness in the proportions shown by Fig. 88 or by the first four columns of Table VIII. The effective colour registered in the brain by light of this wavelength *must*, therefore, be determined by the relative magnitudes of the stimulations which each of these three receptors experiences. Thus we associate the yellow of sodium light ($\lambda = 5.89 \times 10^{-5}$ cm., 5.89×10^2 $\mu\mu$, or 5.89×10^3 A.U.) with relative brightnesses of 241, 160, and 0·01 respectively for the red, green, and blue type receptors. (These values are taken directly from Table VIII.) There is, therefore, no special receptor in the eye for each colour that we see; instead we experience a certain sensation—that we call yellow —when the three receptors are stimulated in the above ratio.

It may seem somewhat strange that such a definite sensation as colour should depend on the mere ratio of the stimulations of only three receptors, but this d. ficulty disappears when we consider the other known peculiarities of the brain.   Thus

we associate, in stereoscopic vision, the relative distances of objects by the slight feeling of strain produced by the two eyes operating at a minute angle with each other. In hearing, we know that—although we can distinguish two notes of nearly the same frequency—we interpret the reception of higher harmonies which are in definite phase relation with the fundamental as changes in the quality of the note heard.

We see from the curves of Fig. 88, or from Table VIII, that the sensation of blue experienced with light of 4.70×10^{-5} cm. must be associated with brightness responses of 6.97, 43.7, and 0.0 for the blue, green, and red receptor respectively; the sensation of green by brightness responses of $\begin{Bmatrix} \text{blue} \\ 0.87 \end{Bmatrix}$, $\begin{Bmatrix} \text{green} \\ 212 \end{Bmatrix}$, and $\begin{Bmatrix} \text{red} \\ 68 \end{Bmatrix}$ respectively; of yellow by $\begin{Bmatrix} \text{blue} \\ 0.07 \end{Bmatrix}$, $\begin{Bmatrix} \text{green} \\ 280 \end{Bmatrix}$, and $\begin{Bmatrix} \text{red} \\ 228 \end{Bmatrix}$ respectively, and of orange by $\begin{Bmatrix} \text{blue} \\ 0.00 \end{Bmatrix}$, $\begin{Bmatrix} \text{green} \\ 61 \end{Bmatrix}$, and $\begin{Bmatrix} \text{red} \\ 208 \end{Bmatrix}$.

The sensation of white is experienced when completely heterogeneous light enters the eye. If this heterogeneous light were to have identically the same energy associated with every wave-length of the spectrum it would appear to be white light with a slight blue tint in it, because normal sunlight contains rather more energy in the yellow and red than in the blue, and we tend to take our standard of white light as being the average noon solar light in our own immediate locality. Since this average noon solar light varies appreciably from the poles to the tropics it must be defined, for the purposes of specification of colour, by international agreement. This was done by the Commission Internationale de l'Éclairage, in 1931 (Cambridge University Press), and three standard whites have been defined. Standard A is typical of the white of artificial light, Standard B of normal solar light in temperate climates, and Standard C of solar light in the tropics,

TABLE VII. STANDARD B WHITE LIGHT

$\lambda \times 10^{-5}$ cm.	3·80	4·00	4·20	4·40	4·60	4·80
Energy	22·4	41·3	63·2	80·8	88·3	95·2
$\lambda \times 10^{-5}$ cm.	5·00	5·20	5·40	5·60	5·80	6·00
Energy	94·2	89·5	96·9	102·8	101·0	98·0
$\lambda \times 10^{-5}$ cm.	6·20	6·40	6·60	6·80	7·00	7·20
Energy	99·7	102·2	105·0	103·9	99·1	92·9

$$10^{-5} \text{ cm.} = 10^2 \mu\mu = 10^3 \text{ A°U.}$$

For this Standard B white light, the total brightness contributions of the three receptors, blue, green, and red, are 1·0%, 53%, and 46% respectively.

Table VIII shows that monochromatic lights with wavelengths between 4·70 × 10⁻⁵ cm. and 5·90 × 10⁻⁵ cm. (i.e. 4·70 and 5·90 × 10² $\mu\mu$ or × 10³ A°U) excite all three receptors and can, therefore, be regarded as mixtures of imaginary 'super' greens or yellows with white. In columns 6, 7, 8, and 9 the total brightness at these wave-lengths is expressed as a mixture of white and of two other receptors. It therefore appears that no normal person sees as vivid a green of wavelength 5·30 × 10⁻⁵ cm., for example, as a person does who is colour-blind to the blue. The only way that a normal non-colour-blind person could be able to envisage such super-pure greens would be to fatigue the eye to the blue receptor by looking at very intense deep violet light and then immediately to look at the spectral green considered. During the short time that the eye remains badly fatigued to the blue, this receptor would produce a very small brightness contribution indeed. This effect is even more marked than would normally be expected, for it has been found that fatiguing the eye in one receptor increases simultaneously the sensitivity of the other two receptors.

The methods by which coloured lights may be specified on this system are best explained by numerical examples.

Table VIII

λ × 10⁻⁵ cm.		BRIGHTNESS OF			TOTAL BRIGHT-NESS	BRIGHTNESS OF				NEGLECTING WHITE		
		BLUE RECEP-TOR	GREEN RECEP-TOR	RED RECEP-TOR		WHITE RECEP-TOR	BLUE RECEP-TOR	GREEN RECEP-TOR	RED RECEP-TOR	% BLUE	% GREEN	% RED
Violet	4·00	2·53	—	—	2·5					100	—	—
	4·10	4·33	—	—	4·3					100	—	—
	4·20	6·14	—	—	6·1					100	—	—
	4·30	9·15	—	—	9·1					100	—	—
	4·40	10·19	3·8	—	14					73·0	27·0	—
	4·50	9·50	8·6	—	18·1					52·4	47·6	—
	4·60	8·42	20·5	—	28·9					29·0	71·0	—
Blue	4·70	6·97	43·7	—	50·7					13·7	86·3	—
	4·80	4·73	66·0	6·3	77	13·8	4·6	58·6	—	7·3	92·7	—
	4·90	2·20	91·5	18·5	112	40·5	1·8	69·7	—	2·52	97·5	—
	5·00	1·23	140	37·4	179	82·2	·46	96	—	·48	99·5	—
Green	5·10	·87	212	68	280	92	—	162	26	—	86·3	13·7
	5·20	·61	276	105	381	65·5	--	241	75	—	76·2	23·8
	5·30	·43	310	136	446	46·4	--	285	115	—	71·1	28·9
	5·40	·29	325	168	493	30·3	—	309	154	—	66·8	33·2
	5·50	·18	330	191	521	19·2	—	320	182	—	63·6	36·4
	5·60	·11	314	210	524	11·1	—	308	205	—	60·0	40·0
Yellow	5·70	·07	280	228	508	7·1	—	276	225	—	55·0	45·0
	5·80	·04	224	234	458	3·0	—	222	233	—	48·6	51·4
	5·90	—	160	241	401					—	40·0	60·6
	6·00	—	106	230	336					--	31·6	68·4
Orange	6·10	—	61	208	269					--	22·7	77·3
	6·20	—	32	169	201					···	15·9	84·1
	6·30	—	15·6	138	124					--	12·6	87·4
	6·40	—	5·4	88	93					--	5·8	94·2
	6·50	—	1·6	53	54					--	2·9	97·1
	6·60	—	—	30·6	30·6					—	--	100
Red	6·70	—	—	18·0	18·0					—	—	100
	6·80	—	—	10·0	10·0					—	—	100
	etc. above 6·80			2·2	2·2							

$$10^{-5} \text{ cm.} = 10^2 \, \mu\mu = 10^3 \text{ A°U.}$$

Table VIII shows the brightness stimulations of the different receptors in arbitrary units when irradiated by a constant energy of monochromatic radiation.

Standard B white has brightness contributions of 1·0%, 53%, and 46% for the blue, green, and red receptors respectively.

§ 79. Interpretation of a Colour whose Blue, Green, and Red Brightness Factors are known

To interpret any light specification on this system, express the values as a mixture of white light and of two other stimulations, making the third brightness value zero.   If we take as

a particular example, light having blue, green, and red brightness factors $\left\{\begin{array}{c}\text{blue}\\21\cdot6\%\end{array}\right\}$, $\left\{\begin{array}{c}\text{green}\\50\cdot8\%\end{array}\right\}$, and $\left\{\begin{array}{c}\text{red}\\27\cdot6\%\end{array}\right\}$, we must make the red receptor zero, since white light is $\left\{\begin{array}{c}\text{blue}\\1\%\end{array}\right\}$, $\left\{\begin{array}{c}\text{green}\\53\%\end{array}\right\}$, and $\left\{\begin{array}{c}\text{red}\\46\%\end{array}\right\}$. Thus in 100 brightness units of the original light $\left\{\begin{array}{c}\text{red}\\27\cdot6\end{array}\right\}$ units goes to form white light in the red receptor, together with $\left\{\begin{array}{c}\text{green}\\53\times\dfrac{27\cdot6}{46}\end{array}\right\}$ in the green receptor, and $\left\{\begin{array}{c}\text{blue}\\1\cdot0\times\dfrac{27\cdot6}{46}\end{array}\right\}$ in the blue receptor. The light is therefore a mixture of 60 brightness units of white spread out among the receptors in the above manner $\left\{\begin{array}{ccc}\text{red}&\text{green}&\text{blue}\\27\cdot6\;+&31\cdot8\;+&0\cdot6\end{array}\right\}$, and of 40 brightness units of light with the values $\left\{\begin{array}{c}\text{red}\\0\end{array}\right\}$, $\left\{\begin{array}{c}\text{green}\\19\end{array}\right\}$, $\left\{\begin{array}{c}\text{blue}\\21\end{array}\right\}$. The mixture is, therefore, one of 60 brightness units of white with 40 of a light having $\left\{\begin{array}{c}\text{blue}\\52\cdot5\%\end{array}\right\}$ and $\left\{\begin{array}{c}\text{green}\\47\cdot5\%\end{array}\right\}$— which is a blue light of $\lambda 4\cdot50\times10^{-5}$ cm.

If this basic wave-length had been between the two wavelengths $\lambda 4\cdot70\times10^{-5}$ and $\lambda 5\cdot90\times10^{-5}$ cm., a part of this 'super-spectral' pure colour would have had associated with it a part of the brightness of the white, since in this region we have regarded the spectral pure colours as mixtures of 'super-pure' colours and white. Thus if the light had been found to be a mixture of, for example, 80 brightness units of white light and 20 of 'super-pure,' $\lambda 5\cdot30$, $\left\{\begin{array}{c}\text{blue}\\0\%\end{array}\right\}$, $\left\{\begin{array}{c}\text{green}\\71\cdot1\%\end{array}\right\}$, $\left\{\begin{array}{c}\text{red}\\28\cdot9\%\end{array}\right\}$, we could have associated

$$\left\{20 \times \frac{\overset{\text{white}}{46\cdot4}}{\underset{285 + 115}{\text{green}\quad\text{red}}}\right\} = 2\cdot32$$

brightness units of the white (see Table VIII, line λ5·30, columns 6, 7, 8, 9) with the spectral colour and would have regarded it as 77·7 brightness units of white light with 22·3 units of λ5·30 × 10⁻⁵ cm.

§ 80. Hue and Saturation

Whatever may be the energy/wave-length distribution of light entering the eye it is detected by the stimulation of the three receptors and can be specified by the corresponding three brightness ratios. Thus any coloured light can in general be regarded as producing a certain number of brightness units of white and a certain number of brightness units in two out of the three receptors. From this point of view any coloured light can be spoken of as being of a certain *saturation*—characterized by the fraction of the brightness associated with the white, and of a certain *hue*—characterized by the wave-length of that monochromatic radiation which —apart from the amount of white present—excites the two activated receptors in the same ratio. This definition of hue naturally needs modification when the dominant receptors are the red and blue, as occurs with magentas and purples.

§ 81. Colour Mixing and Matching

Since any colour, or any shade of any coloured light, can be specified by the relative brightnesses of the three receptors, it must be possible to consider theoretically its production by the addition or subtraction of three—and only three—other colours. There are, then, three unknowns—the necessary energy parts of these three colours—and there are the three equations used for determining the brightness contributions to the three receptors.

Consider, for example, the possibility of exciting the three receptors by x energy parts of $\lambda\,6\cdot80 \times 10^{-5}$ cm. (brightness units $\begin{Bmatrix}\text{blue}\\0\cdot0x\end{Bmatrix}$, $\begin{Bmatrix}\text{green}\\0\cdot0x\end{Bmatrix}$, $\begin{Bmatrix}\text{red}\\10\cdot0x\end{Bmatrix}$), y energy parts of $5\cdot80 \times 10^{-5}$ cm. (brightness units $\begin{Bmatrix}\text{blue}\\0\cdot04y\end{Bmatrix}$ $\begin{Bmatrix}\text{green}\\224\cdot0y\end{Bmatrix}$, $\begin{Bmatrix}\text{red}\\234\cdot0y\end{Bmatrix}$), and z energy parts of $\lambda\,4\cdot40 \times 10^{-5}$ cm. (brightness units $\begin{Bmatrix}\text{blue}\\10\cdot19z\end{Bmatrix}$, $\begin{Bmatrix}\text{green}\\3\cdot8z\end{Bmatrix}$, $\begin{Bmatrix}\text{red}\\0\cdot0z\end{Bmatrix}$), in exactly the same ratio as occurs with $\lambda\,5\cdot00 \times 10^{-5}$ cm. (brightness units $\begin{Bmatrix}\text{blue}\\1\cdot23\end{Bmatrix}$, $\begin{Bmatrix}\text{green}\\140\cdot0\end{Bmatrix}$, $\begin{Bmatrix}\text{red}\\37\cdot4\end{Bmatrix}$).

If the three visual receptors are excited in the same way the colour of the compound mixture will be the same as the colour of the monochromatic comparison light. The condition that they are excited in the same way is clearly that:

$$
\begin{array}{lrcll}
\text{blue} & 0\cdot0x + & 0\cdot04y + 10\cdot19z = & 1\cdot23 \\
\text{green} & 0\cdot0x + & 224\cdot0y + 3\cdot8z & = 140\cdot0 \\
\text{red} & 10\cdot0x + & 234\cdot0y + 0\cdot0z & = 37\cdot4
\end{array}
$$

so that the necessary values of x, y, and z can be found.

It would appear from this example that it must be possible to synthesize theoretically any colour or shade of colour by mixtures of any other three.

Such a synthesis cannot always be performed experimentally because the calculated values of x or y or z are sometimes negative, and we have not a perfect range of light filters by which a particular wave-length can be removed from a beam of light *by any desired amount*.

The usual practical method of *avoiding* the subtraction of light energy from the mixture is by *adding* it instead to the comparison coloured light. Thus supposing we found that, to synthesize a certain brightness of one particular colour, we needed $+ 4\cdot6$ *energy* units of our red standards, $- 2\cdot1$ of our green standard, and $+ 10\cdot3$ of the blue standard, then we should instead obtain a colour match between (our initial

light + 2·1 *energy* units of the green standard) and (4·6 units of the red + 10·3 of the blue standard).

This theoretical treatment of colour mixing may perhaps be made more vivid by considering a possible experimental arrangement as shown in Fig. 89. Light from the various standard colour sources is reflected into the eye by the three *unsilvered* glass mirrors. Each of the sources is provided
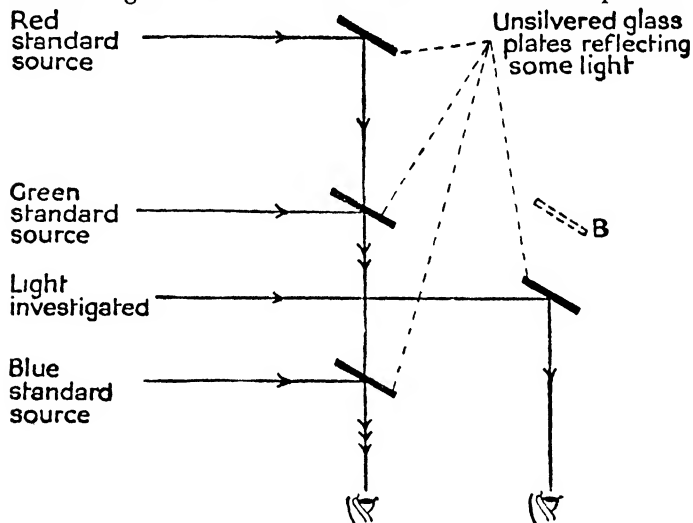

FIG. 89

with a means of altering the light intensity by any desired amount. The light to be investigated falls on to the fourth mirror and so enters the other eye—or another part of the same eye. With this arrangement we are performing a direct addition of the three standards in an endeavour to match in colour and in brightness the light investigated. If, however, it was found necessary, the green reflecting glass plate could be transferred to position B and a match obtained between the red plus the blue, and the green plus the light investigated. Practical instruments for colour matching are naturally much more complicated than this, and the original papers by Guild

or Houstoun should be consulted. Guild recommends for use in such instruments Wratten filter No. 71 (central band $6·3 \times 10^{-5}$ cm.), No. 62 (central band $5·37 \times 10^{-5}$ cm.), and No. 49B (central band $4·5 \times 10^{-5}$ cm.).

§ 82. Colour Mixing to produce White Light

It is found experimentally that white light can be produced by mixing any monochromatic radiation between $6·55 \times 10^{-5}$ cm. and $5·64 \times 10^{-5}$ cm. with appropriate amounts of *one* other monochromatic radiation whose wave-length is always between $4·96$ and $4·29 \times 10^{-5}$ cm. Thus, for example, if

$1·32$ *energy parts* of $\lambda\ 5·74 \times 10^{-5}$ cm. (brightness $\begin{Bmatrix} \text{blue} \\ 0·55 \times 1·32 \end{Bmatrix}$, $\begin{Bmatrix} \text{green} \\ 252 \times 1·32 \end{Bmatrix}$, $\begin{Bmatrix} \text{red} \\ 234 \times 1·32 \end{Bmatrix}$) is mixed with 1 energy part of $\lambda\ 4·72 \times 10^{-5}$ cm. (brightness $\begin{Bmatrix} \text{blue} \\ 6·6 \end{Bmatrix}$, $\begin{Bmatrix} \text{green} \\ 48 \end{Bmatrix}$, $\begin{Bmatrix} \text{red} \\ 1·5 \end{Bmatrix}$), the total brightness values will be $\begin{Bmatrix} \text{blue} \\ 7·3 \end{Bmatrix}$, $\begin{Bmatrix} \text{green} \\ 381 \end{Bmatrix}$, and $\begin{Bmatrix} \text{red} \\ 310 \end{Bmatrix}$, or $\begin{Bmatrix} \text{blue} \\ 1·0\% \end{Bmatrix}$, $\begin{Bmatrix} \text{green} \\ 55\% \end{Bmatrix}$, and $\begin{Bmatrix} \text{red} \\ 44\% \end{Bmatrix}$, which are identical with the brightness values for white light. Pairs of colours which yield white in this way are called complementary colours.

Other complementary monochromatic colours are $\lambda\ 6·50$ and $4·96 \times 10^{-5}$ cm., $\lambda\ 5·80$ and $4·82 \times 10^{-5}$ cm., and $\lambda\ 5·70$ and $4·43 \times 10^{-5}$ cm.

The production of white light by mixing energy portions x, y, and z of three monochromatic radiations is naturally only a special case of the general treatment given on page 160. As an illustration, consider the mixture of x energy units of $\lambda\ 7·00 \times 10^{-5}$ cm. (brightness $\begin{Bmatrix} \text{blue} \\ 0·0 \end{Bmatrix}$, $\begin{Bmatrix} \text{green} \\ 0·0 \end{Bmatrix}$, $\begin{Bmatrix} \text{red} \\ 2·21x \end{Bmatrix}$) with y units of $\lambda\ 5·46 \times 10^{-5}$ cm. (brightness $\begin{Bmatrix} \text{blue} \\ 0·22y \end{Bmatrix}$, $\begin{Bmatrix} \text{green} \\ 330y \end{Bmatrix}$, $\begin{Bmatrix} \text{red} \\ 191y \end{Bmatrix}$), and with z units of $\lambda\ 4·358 \times 10^{-5}$ cm.

(brightness $\begin{Bmatrix} \text{blue} \\ 9.8z \end{Bmatrix}$, $\begin{Bmatrix} \text{green} \\ 0.0 \end{Bmatrix}$, $\begin{Bmatrix} \text{red} \\ 0.0 \end{Bmatrix}$) to make white light with

brightness values $\begin{Bmatrix} \text{blue} \\ 1 \end{Bmatrix}$, $\begin{Bmatrix} \text{green} \\ 53 \end{Bmatrix}$, and $\begin{Bmatrix} \text{red} \\ 46 \end{Bmatrix}$. Then:

blue receptor $1 = 0.0 \quad + \quad 0.22y + 9.8z$
green receptor $53 = 0.0 \quad + 330y \quad + 0.0$
red receptor $46 = 2.21x + 191y \quad + 0.0$

and so $y = 0.161$, $x = 6.92$, and $z = 0.1$.

This example shows that the sensation of white may be produced by mixing suitable (positive) quantities of mono-chromatic red, green, and blue lights. Had the green and the blue been initially mixed together, the white sensation would have been produced by the addition of the red alone. The red would then be termed the complementary colour to the mixture.

In a similar way complementary colours can be found with mixtures which are not monochromatic. This is easily under-stood, since we know that any compound mixture of a large number of different monochromatic waves will stimulate the three receptors in some particular ratio and can therefore be regarded as a mixture of white with one dominant spectral (or purple) colour.

§ 83. Other Systems of Colour Specification

In view of the experimental fact (explained above) that a perfect match for any colour or shade of colour can be produced by the addition or subtraction of any three other colours, it is naturally possible to specify a colour by the number of units of those three which are needed to produce the perfect match. Each of these possible specification systems will be equally logical and correct, so that in each experimental apparatus a different system will be employed and different numerical results will be obtained.

The outline given above (pages 159–163) shows, however, that the brightness contributions of the three receptors can be determined for unit energy of the three standards of any

specification system. If, therefore, we find experimentally the energy proportions of these three standards necessary to match a given colour shade, we can find the brightness contributions of the three receptors when stimulated by this colour shade.

This same colour shade could now be synthesized from three *different* standard colours *belonging to a different reference system,* and—once the brightness values of unit energies of these three standards have been found—the relative energy proportions of these three standards required to match the colour can be calculated. It is, therefore, possible to convert energy (or brightness) units from one system to another.

It was the normal practice at the National Physical Laboratory to use as standards the three wave-lengths $\lambda\,7 \cdot 00 \times 10^{-5}$ cm., $\lambda\,5 \cdot 461 \times 10^{-5}$ cm., and $\lambda\,4 \cdot 358 \times 10^{-5}$ cm. It was also usual to express the relative proportions of these quantities, not in terms of energy units, or in terms of brightness units, but in terms of arbitrary units of such a size that equal numbers of red, green, and blue make white light. These units are called trichromatic units and can be connected with energy units by means of the calculation on page 162, in which it was seen that $6 \cdot 9$ energy units of red, $0 \cdot 161$ of green, and $0 \cdot 1$ of blue were needed to make white light. A specification table of the spectral colours on this system is given on pages 20 and 21 of the report of the Commission Internationale de l'Éclairage of 1931 (Cambridge University Press).

In these trichromatic units, a colour λ may be specified:

$$Q_\lambda = r\mathrm{R} + g\mathrm{G} + b\mathrm{B},$$

where Q_λ, R, G, and B stand for the trichromatic units of the four lights respectively, and where r, g, and b represent the number of these units required to produce 1 unit of the colour concerned. The unit of the colour Q_λ is adjusted to be of such a size that $r + g + b = 1$.

By international agreement, the Commission decided to adopt new standards for the specification of colours, and to convert the experimental values of the proportions of the

different standard colours required to match any given colour shade into these international standards. Table IX shows the numerical values of the specifications of different spectral colours.

TABLE IX. TRICHROMATIC COEFFICIENTS ON INTERNATIONAL SPECIFICATION, 1931

$\lambda \times 10^{-5}$ cm. $\times 10^{+2} \mu\mu$ or 10^{+3} A.U.	3·80	4·00	4·20	4·40	4·60	4·80	5·00	5·20
X	0·174	0·173	0·171	0·164	0·144	0·091	0·008	0·074
Y	·005	·005	·005	·011	·030	·137	·538	·834
Z	·821	·822	·823	·825	·826	·762	·453	·092

<------------------ Violet ------------------>   ↑ Blue   ↑ Green

$\lambda \times 10^{-5}$ cm. $\times 10^{+2} \mu\mu$ or 10^{+3} A.U.	5·40	5·60	5·80	6·00	6·20	6·40	6·60	6·80
X	0·230	0·373	0·512	0·627	0·692	0·719	0·730	0·734
Y	·754	·624	·487	·372	·308	·281	·270	·266
Z	·016	·002	·001	·000	·000	·000	·000	·000

↑ Yellow   ↑ Orange   <------------ Red ------------>

The units of this specification are such that white (this white has equal energy in every wave-length) has the values $\begin{Bmatrix} x & y & z \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{Bmatrix}$. Table IX and the specification value of white should enable rough visual estimates to be made of the international specification values of any coloured light.

It is rather too difficult to give here a scientific definition of these international units x, y, and z. It must suffice, therefore, to mention that the units have been chosen to permit colorimetric calculations to be made with the least arithmetical work, and have not been specially arranged to be easy to understand theoretically. In addition, it will be noted that the specification values of all spectral (and other)

colours have positive values on this system. The three standards are imaginary quantities and may be thought of as super-colours which are only found mixed together in nature. The more of X there is, the redder the colour, the more of Y the greener, and the more of Z the bluer.

One of the reasons for using such peculiar standards is that they have been chosen so that both the red and the blue standards affect the colour of the resulting light without contributing in any way to the brightness. The brightness is, therefore, solely controlled by the green standard—which reduces to $\frac{1}{3}$ the arithmetical labour of this part of brightness calculations. (The Y values of Table IX do not represent the apparent visual brightness of equal energies of the different colours, since they are expressed in international trichromatic units.)

§ 84. Warnings

A number of points dealt with in this chapter may cause difficulty, and it seems convenient at this stage to examine them further.

(a) It is fundamentally more scientific to specify colours and to discuss the relations between them from the standpoint of the colour-mixing of three coloured lights than from that of the receptor theory. The possibility of obtaining colour matches from three colours is known experimentally, while the three-receptor theory has been built up to explain these facts. Since the brightness/wave-length curves of the three receptors were produced essentially by a combination of the colour-mixture data with information concerning the colour-blind, the accuracy of these curves—and of conclusions based on them—must necessarily be less than the accuracy of the colour-matching experiments. The whole phenomenon of colour mixing is described here in terms of the sensation curves of the three receptors, because the theory is simpler to understand from this point of view, and because it undoubtedly approximates closely to the actual behaviour of the eye.

(*b*) There is still some experimental uncertainty regarding the relative constancy of the three brightness/wave-length curves at low intensities. It is certain that at very low intensities the total sensitivity curves of the cones and rods alter (Purkinje effect), and the only doubt concerns vision in which the image is formed on the yellow spot of the retina— where there are no rods and only cones. It is, however, certain that over the very wide range of normal experimentation the sensitivity curves of the eye are constant, and only vary appreciably at very high or at very low intensities.

(*c*) Slight variations occur between the sensitivity/wave-length curves of different observers; for this reason the proportions of three standards required to match any other light will vary slightly from observer to observer. The corresponding colour specifications will therefore also vary *slightly*, and to avoid this difficulty a standard visibility curve has been agreed upon and a standard observer defined by the different proportions of wave-lengths 7·00, 5·461, 4·358 × 10^{-5} cm., which are required to match different monochromatic radiations (Commission Internationale de l'Éclairage, 1931). This standard observer has eye characteristics which are certainly very close indeed to those of the large majority of people.

(*d*) Men who are totally colour-blind in one of their visual receptors have colour sensations and can perform colour-matching experiments. Since they have only two conditions to satisfy they need only two specification values, and two, not three, lights to match any other coloured light.

(*e*) In many books visual sensation curves are given which look very different from those of Fig. 88 and Table VIII. This is because each of the curves has been scaled arbitrarily by different amounts in such a way that the three (red, green, and blue) ordinates of white light are identical. The ordinates do not, therefore, really represent the separate receptor contributions to the total brightness.

(*f*) The colour triangle is generally mentioned and explained in the usual books on light. We have shown that any colour may be specified by the relative proportion, in any convenient

units, of three standard colours, and that unit quantity of it may be expressed by the number r of units of red light, the number g of units of green light, and the number b of units of blue light, the total $(r + g + b)$ being equal to unity. A colour may then be represented by a point in an equilateral triangle, the numerical values of r, g, and b being proportional
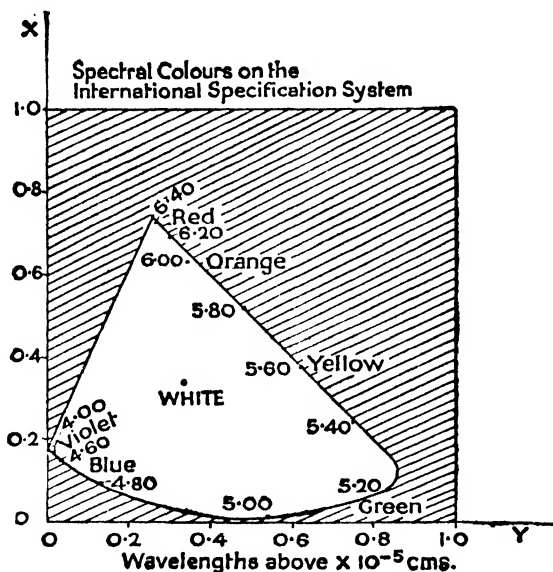

FIG. 90

to the perpendicular distances of this point from the three sides of the triangle. This follows because in an equilateral triangle the sum of these perpendiculars is constant whatever may be the position of the point. Finally, if the magnitudes of the units of the three standard lights have been chosen in the conventional manner so that white light is represented by $\frac{1}{3}$ of each unit, the centre of the triangle will correspond to white light.

A more common modern method of representation is to

employ the usual Cartesian right-angle axes, to record the numerical values of r and g by the two co-ordinates, and to calculate the b value from the equation $r + g + b = 1$. If this is done when the colours are specified on the new international units, all colours possible will be found to lie inside the area shown in Fig. 90.

(g) A further point of difficulty sometimes arises over the meaning of the term 'primary colour.' This term is used commonly by three different groups of people with three entirely different meanings, and therefore occasions much confusion. Physicists usually employ it to denote the three standard colours that they are using at the moment to match the colour investigated; if they change these three wavelengths they say that they have changed their primary colours. To the physicist, therefore, any colour may be primary. Artists usually say that red, blue, and yellow are primary colours, because they find that they can obtain *paint* of any other hue by appropriate mixing of these three. To be really precise they should say that these are primary pigments and not primary colours. Psychologists usually say that red, yellow, green, blue, and white are primary colours, because most people describe colours in terms of these five. In opposition to this phraseology it is certain that the eye does not contain yellow visual receptors, for Hecht has shown that we see yellow if one eye receives red light and *if the same part* of the other eye is irradiated with green. In a similar manner the sensation of white may be simulated from yellow in one eye and blue in the other.

§ 85. References

Since very little information on modern developments of colorimetry has appeared as yet even in the advanced textbooks, readers desiring further information must consult the original papers. The following selection may be of use and will provide any further references desired:

Jour. Frank. Inst., clxxx, p. 409 (1915). Ives discusses here the relation between different spectral primaries in colour mixing and obtains sensation curves from König's data.

Jour. Opt. Soc. Amer. and Rev. Sci. Inst., vol. 6, p. 549 (1922). Weaver brings the results of all experiments on colour mixing to a common basis and shows how to obtain the sensation curves which were adopted by the Optical Society of America.

Trans. Opt. Soc. Amer., vol. 26, p. 95 (1924–5). Guild simplifies the transformations involved in changing from one trichromatic system to another.

Report of Commission Internationale de l'Éclairage, 1931 (Camb. Univ. Press). Contains tables of colour mixing, visibility curves of a standard observer, and the international specification of the spectral colours.

Trans. Roy. Soc. A., vol. 230 (1931). Guild gives a comprehensive article on the colorimetric properties of the spectrum.

Trans. Opt. Soc., 1931–2. Contains papers by Guild and Smith on the international colorimetric specifications and their uses.

Report of Discussion on Vision (the Physical Society, 1932). Contains many papers on problems of vision and colorimetry.

Jour. Opt. Soc. Amer., vol. 23, p. 359 (1933). Judd gives a comprehensive article on colorimetry.

§ 86. Light Filters

Coloured glasses, coloured liquid solutions, or dyed gelatine films form the usual light filters. The pigment possesses the power of absorbing completely certain wave-length regions of the spectrum, so that, if white light is incident, the transmitted light is coloured complementary to the spectral parts which are absorbed. The energy which is thus absorbed produces heat, so that care must be taken when using gelatine filters to prevent the gelatine melting. Much of the light reflected from a filter has penetrated a finite depth and so will also have lost all those radiations whose wave-lengths were within the absorption band. The reflected light will, therefore, have almost the same colour as the transmitted light.

If light traverses a series of different filters, the first will remove completely all those components of the light which lie in its own spectral absorption region, while the others, in turn, will remove any of the components which are still left and which lie in their spectral absorption region. The colour of the transmitted light (if any) will, therefore, be due to those wave-lengths of the original light which were absorbed by none. If the absorption bands overlap completely, there will be no light penetrating the filter.

§ 87. Paint

We see coloured objects by the light which they reflect or diffuse into the eye. The spectral distribution of this diffused
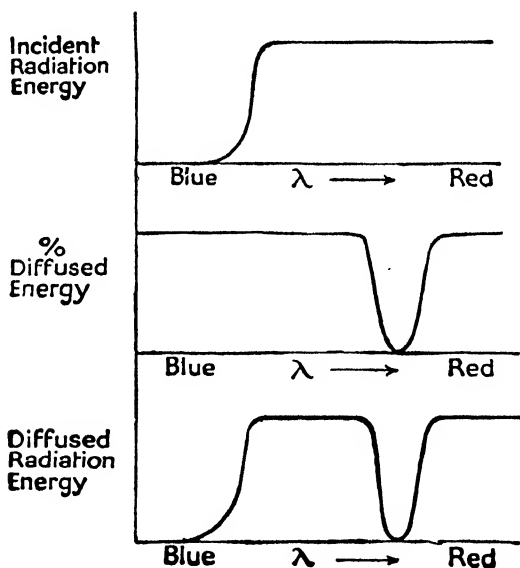
FIG. 91

light depends on that of the primary light, and on the variation with wave-length of the percentage of energy reflected or absorbed by the coloured object. Fig. 91 illustrates

how the spectral composition of light may be altered by reflection from an absorbing substance.

A perfect red paint would reflect radiations between 6·20 and 6·80 times 10^{-5} cm. only, and would absorb completely the whole of the remainder of the energy. Such a paint would look red whatever might be the spectral distribution of the incident light providing it did contain some red component. If the incident light was completely deficient in this wave-length region, the paint would look black, because no light would be diffused or reflected from it.

From their very nature most paints are not 'pure,' and reflect an appreciable fraction of energy at all wave-lengths.

FIG. 92

They are coloured because they reflect certain components of the spectrum very efficiently indeed. Such paints would be described as 'unsaturated,' since the diffused light could be regarded as a mixture of white and coloured lights. The colour changes shown by such paints in light of differing spectral distribution are always interesting. These changes can be guessed at in a general way, if the spectral composition of the light and the reflection factors of the surface are known, or, of course, they could be calculated accurately by extensions of the methods already discussed in colour mixing.

We may now consider the colours shown in white light, or, for preference, the reflection-factor wave-length graphs, of mixtures of different pigments. Such paint mixtures must consist of a complete mixture of tiny grains of the absorbing pigments as shown in Fig. 92. Let us first suppose that the pigment is such that light never penetrates completely through a grain, that the reflection occurs at the surface of the grain, and that the two pigments have, for example, absorption bands covering the whole of the spectrum except

for the red and the blue respectively. One-half of the incident rays will encounter the red grains, and one-half the blue, so that the reflected, or diffused, light will be a mixture of one-half the normal brightness of red and of blue. This example shows that a mixture of two such pigments produces colours which are less saturated and which approximate to white, and that the total energy reflected from the paint will be much less than 100% and will be smaller the more spectrally pure were the initial paints. This is illustrated by Fig. 93, in which the spectral distribution curve of the secondary light from the above paint mixture is compared with that of the incident radiation. We shall show later that a white of low relative brightness is interpreted by the eye as grey, so that the above mixture

Fig. 93

will act as a dull or greyish purple. The most vivid purple would be one which reflected 100% each of the blue and of the red components, and not 50%, as in this example.
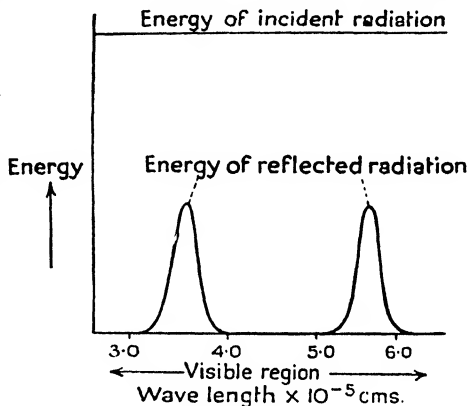
The artist finds it possible to produce colours of any hue by mixing his three primary paints, red, yellow, and green. He generally purchases more colours than these, because his synthetic colour is somewhat unsaturated, and because it is much less bright than the maximum possible.

Another type of paint exists, which does not contain strongly reflecting coloured grains, but which instead employs a white base of some kind, and highly coloured, transparent dyes. This may be exemplified by the application of mixed coloured inks to a white paper surface. The mixture laws for such

paints are essentially those governing the transmission of light through filters, and are therefore entirely different from those of the first type of paint. With the former a mixture of equal parts red and blue means that one-half as much blue is reflected as with the blue alone, and one-half as much red as would have been obtained from the red alone. The resulting colour is then dull purple. With the second type of paint the light has to penetrate the two filters in turn; the blue filter will absorb everything with the exception of the blue rays, while the red filter will in turn absorb these rays also, since it transmits red only. The resulting colour will then be black, if the dyes are perfect.

A modification of this second type of paint will include the dyes, and, in itself, a finely powdered white base. The colour of mixtures of such paints will then be that produced by radiation transmitted through both dye filters, but will be less saturated because some of the white base grains will be so near the surface that the light filters would be too thin to be effective.

§ 88. Paint—Browns and Greys

Among coloured materials there exist certain drab colours —or broken tones—such as browns, olive greens, greys, etc., which do not occur in the spectrum and which cannot be produced by merely mixing coloured lights. Such colours owe their peculiar appearance only to their relatively low brightness *seen in contrast to surfaces of similar hue but of much greater brightness*. Thus the total fraction of the incident light reflected by a bright impure yellow is, for example, about 70%, while that from a similar brown is only 6% or 7%.

It is easy to show experimentally that greyness is a simple contrast effect of white by observing a beam of strong sunlight entering a window to form its sharp edge on a black sheet or cloth. Place a grey paper in the sunlight and a white paper in the shadow, with their junction line exactly

along the shadow boundary. The white paper will then seem grey, and the grey, white.

It follows from this that the paint specification should include not merely a graph of reflection intensity against wave-length, but should also have specified on the graph the magnitude of the ordinates, so that the total reflection factor may be determined.

A much less informative method of paint specification is to give the total brightness factor and the international specification of the colour of the light reflected. This is less informative because the colour of the reflected light depends on the spectral distribution of the primary, and so, if this is done, one of the three international standard illuminations A, B, or C should be employed.

CHAPTER VIII

INTERFERENCE

§ 89. Introduction. § 90. Methods of producing Two Sources in Phase. § 91. Interference of Light reflected from Thin Films. § 92. The Colours of Thin Films. § 93. The Position of Newton's Fringes. § 94. The Experimental Investigation of Thin Films. § 95. Circular Fringes. § 96. Newton's Fringes. § 97. The Effect of silvering Thin Films. § 98. The Michelson Interferometer and the Wave-length of Light.

§ 89. Introduction

The whole theory of interference depends on the assumption that the total amplitude at a point, or the total force experienced by an electron, is the vector sum of the separate amplitudes or forces produced by all the different light waves which may be present at the point. Thus if two wave motions,

$$y_1 = a_1 \cos pt,$$
and
$$y_2 = a_2 \cos (pt + \delta),$$

are present simultaneously at the same point and are operating in the same direction, then the resultant amplitude is the sum of the two separate contributions and

$$y = y_1 + y_2 = a_1 \cos pt + a_2 \cos (pt + \delta).$$

Now we will show later (page 192) that this is identical with another simple harmonic oscillation of the same frequency, and

$$y = \left(\sqrt{a_1{}^2 + a_2{}^2 - 2a_1a_2 \cos \delta} \right) \cos (pt + \beta),$$

where β is a phase angle.

If $\delta = \pi$, or 3π, etc., the *maximum* amplitude will be $(a_1 + a_2)$, while if $\delta = 2\pi$ it would be $(a_1 - a_2)$. This shows that, if the separate component amplitudes were equal, the resulting intensity or energy ($\propto$ (amplitude)^2) would be four

times the normal value if the two beams were in phase, and would be zero if the beams were exactly out of phase.

The simplest example, in theory, of interference refers to two sources of the same frequency and of the same phase and which are separated by a very small distance b (Fig. 94).

FIG. 94
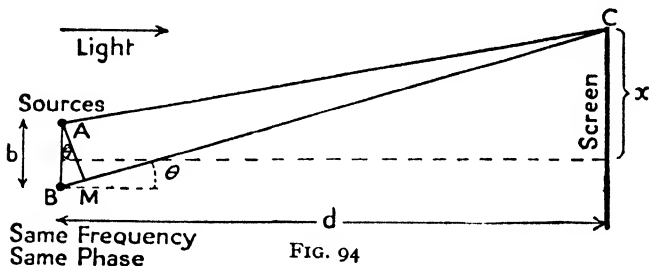
The light reaching the point C from the two sources A and B must be out of phase with each other by an amount governed by the difference in distance the two waves have travelled (BC − AC). There will be a maximum of intensity at C if

$$(BC - AC) = n\lambda,$$

i.e. (to a high degree of approximation) when BM = $n\lambda$.

$$\therefore b \sin \theta = n\lambda;$$

but tan θ is really equal to $\left(\dfrac{x + \dfrac{b}{2}}{d}\right)$, which is $\dfrac{x}{d}$ if b is small

compared with x; thus $b\left(\dfrac{x}{d}\right) = n\lambda,$

and $$x = n\left(\dfrac{d}{b}\right)\lambda.$$

Thus at regular distances from the centre there will be light bands alternating with regions of darkness for which

$$x = (n + \tfrac{1}{2})\left(\dfrac{d}{b}\right)\lambda. \checkmark$$

If d is 100 cm. the bands will be separated by distances of

$$\left(\dfrac{5 \times 10^{-3}}{b}\right) \text{ cm.} = \left(\dfrac{1}{20}\right)\left(\dfrac{1}{b}\right) \text{ mm.}$$

G

·It therefore follows that if the fringes are to be seen without magnifying arrangements, b must be of the order of $\frac{1}{20}$ cm.

§ 90. Methods of producing Two Sources in Phase

We must emphasize that the interference phenomena will be observed only if the two sources have *always* the same phase or have always the same constant phase difference. Normal monochromatic light consists of trains of waves stopping suddenly when the parent atom makes an atomic collision, and restarting with an entirely arbitrary phase relation when
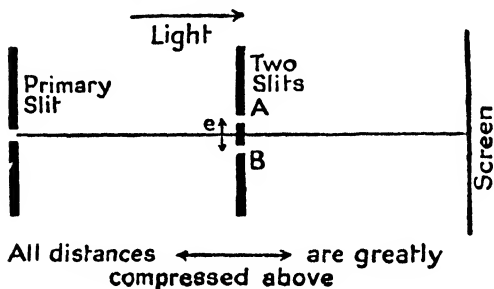
Light

Primary Slit

Two Slits A

e

B

Screen

All distances ←——→ are greatly compressed above

FIG. 95

the parent atom or molecule recovers from the collision. ·This process occurs irregularly thousands of times a second, so that if at one instant the phase difference between two separate sources was such that a bright band was produced at one particular point on the screen, then a thousandth of a second later the phase difference and the position of the point of maximum intensity would have changed. The only possibility, therefore, of observing the fringes is to arrange the experiment so that the light from the two sources comes originally from the same primary source. If the phase difference of the first source changes arbitrarily as a result of an atomic or molecular collision, that of the second source will also change by the same amount, so that the *difference* between the two phases will be constant.

The experimental arrangement of Thomas Young is shown in Fig. 95. The two slits A and B must always be in the same

phase since they are both illuminated by the same primary source.

If the pin-holes of the theory are replaced by the slits in an endeavour to obtain more light energy, the slits must be *accurately parallel* to each other.

Fresnel mirrors. This method employs two mirrors placed at a very small angle to produce two virtual images of the primary source. The student should calculate the distance between the two images A and B in terms of the angle *â* between the two mirrors, with the help of Fig. 96. Clearly the interference pattern can only be observed in the shaded area where the light is received from *the two* mirrors.
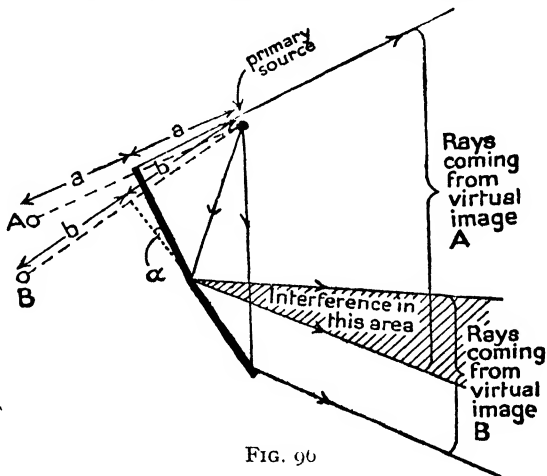
FIG. 96
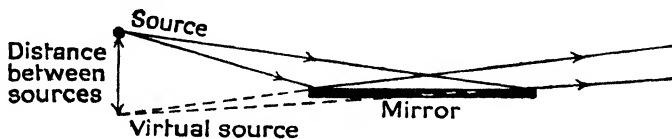
FIG. 97

Lloyd's Mirror. The mirror should be formed by silvering the front surface of a sheet of optically worked glass.

The interference then occurs in the region traversed by the direct rays from the source and by the rays reflected by the

mirror (Fig. 97). This region of interference is clearly very small unless the mirror is exceedingly long or unless the source is very near to it. The importance of the Lloyd mirror is that, if the distance $\left(\dfrac{b}{2}\right)$ of the source above the plane of the mirror is altered proportionally to the wave-length, it is possible to obtain achromatic (black and white) fringes. This alteration of the distance $\left(\dfrac{b}{2}\right)$ with the wave-length is effected by the use of a grating and a normal spectrometer arrangement of lenses.

Fresnel bi-prism. Fresnel employed two prisms of exceedingly small angle (Fig. 98). The angle of deviation produced by a small angle prism is constant and is $(\mu-1)a$. The rays, after refraction, therefore come from the two virtual images shown and

$$\frac{b}{2} = l \text{ (angle of deviation)} = l(\mu - 1)a.$$
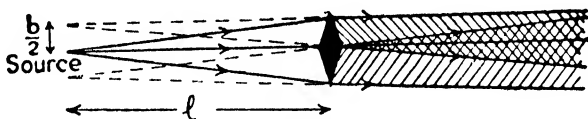
FIG. 98

Once again the interference can occur only in the shaded region which is irradiated by light from *the two* images.

Additional experimental arrangements are the Billet split lens and the bi-plate. These are explained by Figs. 99 and 100.
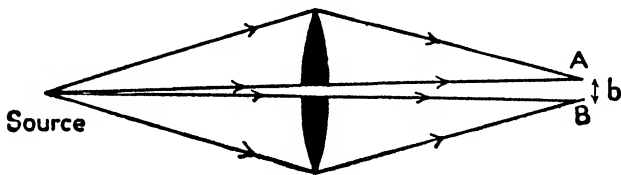
FIG. 99
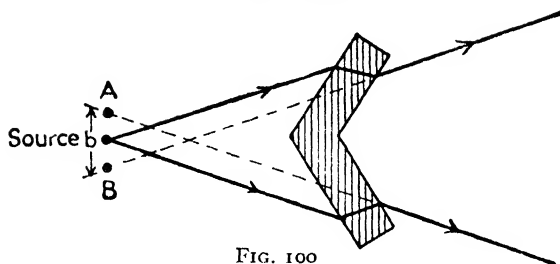
FIG. 100

§ 91. Interference of Light reflected from Thin Films

The most common practical example of the interference of
light occurs when light is reflected from thin films. The
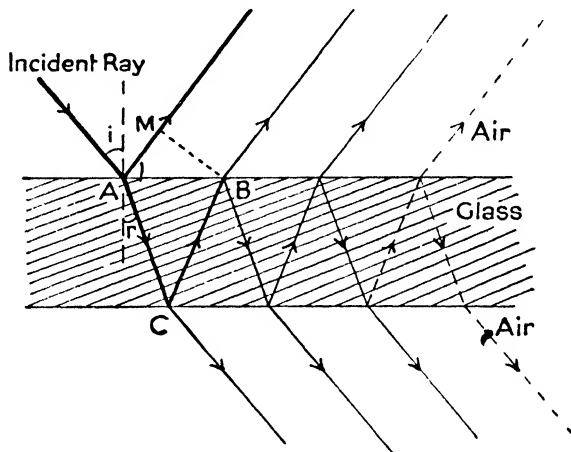
FIG. 101

interference occurs between the light partially reflected from
the top and from the bottom surface of the film. Consider
a ray making an angle of incidence i with a thin film of
refractive index μ. Rays of gradually diminishing intensity
will be reflected and transmitted as shown in Fig. 101. The
path difference between each of these reflected (or trans-
mitted) rays will be all equal and will be $\mu(AC + CB) - AM$.

Now $\qquad$ AM $=$ (AB)(sin i) $=$ ($2e$ tan r)(μ sin r)$,$

and $\qquad$ AC $=$ CB $= \dfrac{e}{\cos r}.$

$\therefore$ path difference $= \dfrac{2\mu e}{\cos r} - 2\mu e$ tan r sin r

$$= \dfrac{2\mu e}{\cos r}(1 - \sin^2 r) = 2\mu e \cos r.$$

The angle involved is the angle inside, and not outside, the film; it is, therefore, the angle which is actually measured only when the film is of air and has a refractive index of
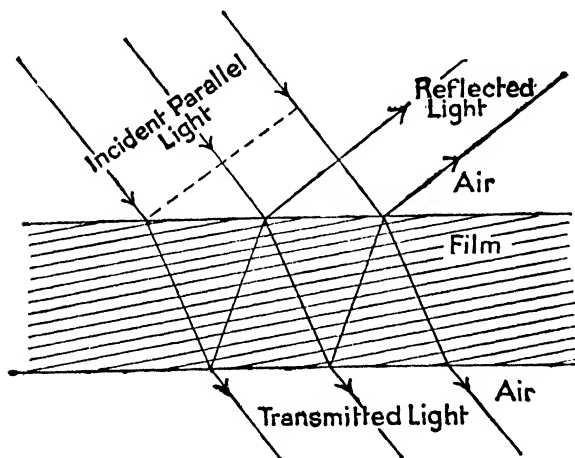
FIG. 102

unity. The interference occurs when all the emerging rays have been brought together at the focus by a lens, and there will be a large intensity at one point in the focal plane of the lens (or on the retina, if an eye accommodated for infinity is employed directly), if the path difference between the primary and the reflected parts of the ray is suitable. Correspondingly, a minimum of intensity will occur at slightly different values of r and of the path difference. A consideration of the in-

tensity of the *reflected* component when the film thickness is reduced to zero, shows that different phase changes must occur at the reflecting surfaces of such a nature that the actual path difference between the two ray parts has the value above plus $\dfrac{\lambda}{2}$ and is therefore:

$$2\mu e \cos r + \frac{\lambda}{2}.$$

Notice that the interference occurs between parts *of the same ray* after the parts have been brought to a focus. It is, therefore, *not necessary* to draw a diagram, as many books do, which assumes that the incident light is a parallel beam. A typically unsatisfactory diagram of this nature is shown in Fig. 102.

§ 92. The Colours of Thin Films

Very interesting and important colour effects occur when white light is reflected from thin films such as thin glass films, or oil films on puddles, etc.

The rays coming to the eye proceed, in general, from different thicknesses (e) of film (Fig. 103), and certainly have different values of $\cos r$; thus the path difference for light entering the eye varies over the whole film. There will be a maximum of intensity recorded at all angles and at all wave-lengths for which the total path difference
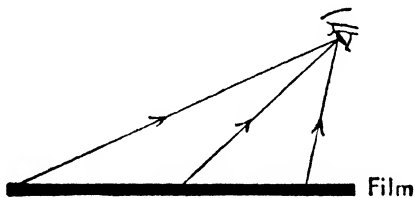
FIG. 103

$$2\mu e \cos r + \frac{\lambda}{2} = n\lambda,$$

and a minimum when

$$\left(2\mu e \cos r + \frac{\lambda}{2}\right) = (n + \tfrac{1}{2})\lambda.$$

The actual colour shown by any part of the film thus depends

on the wave-lengths of those radiations which are strongly reflected. The matter may be made clear by some examples.

Suppose $\left(2\mu e \cos r + \dfrac{\lambda}{2}\right) = 4 \times 10^{-5}$ cm., then the intensity/wave-length graph of Fig. 104 shows that this reflected light would be blue.
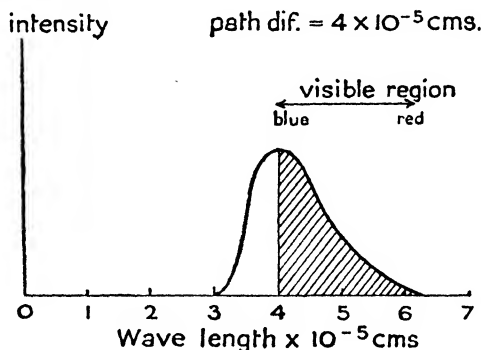
FIG. 104

If the path difference were 8×10^{-5} cm. the intensity maxima would be at $\dfrac{8 \times 10^{-5} \text{ cm.}}{1}$, $\dfrac{8 \times 10^{-5} \text{ cm.}}{2}$, $\dfrac{8 \times 10^{-5} \text{ cm.}}{3}$, etc., and the minima at $\dfrac{8 \times 10^{-5} \text{ cm.}}{1\frac{1}{2}}$, etc. The colour seen would be a somewhat impure purple or mixture of red and blue. (Fig. 105.)
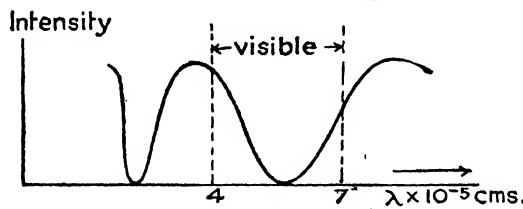
FIG. 105

As a final example consider the intensity reflected from a film when the path difference is large, say 20×10^{-5} cm. (Fig. 106). The maxima will occur at wavelengths such that $n\lambda = 20 \times 10^{-5}$ cm., where n is integral, i.e. at $\lambda = 20 \times 10^{-5}$ cm., $\lambda = 10 \times 10^{-5}$ cm., $\lambda = 6.66 \times 10^{-5}$ cm.,

etc. Similarly minima occur when $n =$ (an integral $+ \frac{1}{2}$). Fig. 106 illustrates that the rays reflected in the visible region would be so widely distributed throughout the spectrum that they would form at least an approximation to white light and that this approximation would be more and more exact as the path difference was increased.

FIG. 106
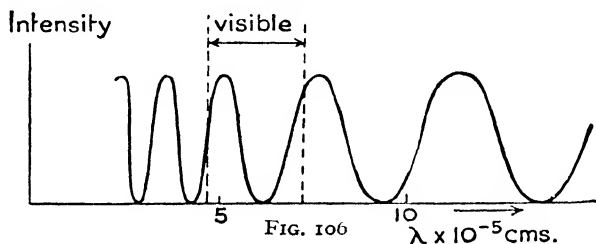
Although we have explained that the *eye* could not detect colour in the light reflected from a thick film with a large path difference, a spectrometer could naturally always detect dark bands occurring at the wave-lengths for which n is (integral $+ \frac{1}{2}$).

§ 93. **The Position of Newton's Fringes**

The interference fringes observed in monochromatic light from thin wedge films are very interesting and important. Consider one single ray incident on one single point D on a wedge film (Fig. 107). The two most important reflected rays are shown meeting at the point P, the path difference between them being (DC + CA − DB), which we can see (and can prove) is, to a high degree of approximation, the value obtained with a parallel film ($2\mu e \cos r$).

The point P where the rays intersect will be *the position* at which the interference fringes will be formed, and screens must be placed at P to obtain a clear image, microscopes must be focused on P, or optical systems arranged to project images of objects placed at P; the interference occurs only at the point P, where all the different parts of the *same* ray

*G

intersect. Many books are far from clear on this point and suggest that the interference occurs anywhere, and is between one part of one incident ray and another part of another ray. This can occur only if the original source of light is a point from which all of the rays diverge with definite phase relations with each other.

The distance AP, or BP, or DP (they are all approximately the same) of this point of intersection of the rays (i.e. of the point where the interference fringes are formed) may be found
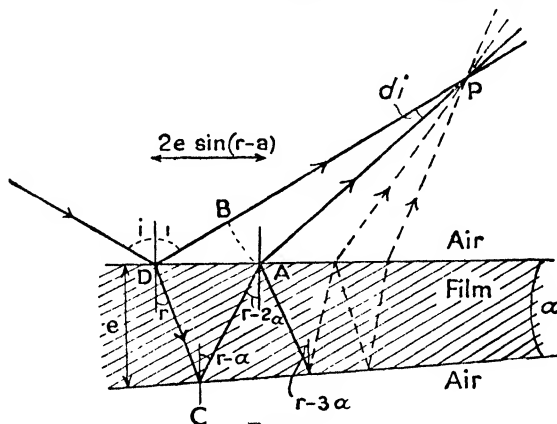
FIG. 107

by calculating first DA in terms of e, then BA, and finally by noting that $AP = \dfrac{BA}{\overset{\wedge}{di}}$ This angle $\overset{\wedge}{di}$ can be found because $\sin i = \mu \sin r$,

$$di \cos i = \mu dr \cos r,$$

and the figure shows that $\overset{\wedge}{dr}$ has the small value $2a$. This distance of the interference fringes from the film is shown finally to be

$$\frac{e}{\mu a} \sin r \frac{\cos^2 i}{\cos^2 r}.$$

Now $\dfrac{e}{a}$ is the distance along the wedge to the apex (or edge)

from the neighbourhood concerned. If this distance is x, the distance of the interference fringes from the film will thus be of the order of $(x \sin r)$. It will therefore generally be very small in all cases except those in which the rays are incident at a large angle with the normal, or in which the two sides of the film are accurately parallel, when a telescope set for infinity must be employed.

§ 94. The Experimental Investigation of Thin Films

The normal conditions under which thin films are studied are shown in Figs. 108 (a), (b), (c), (d), (e).
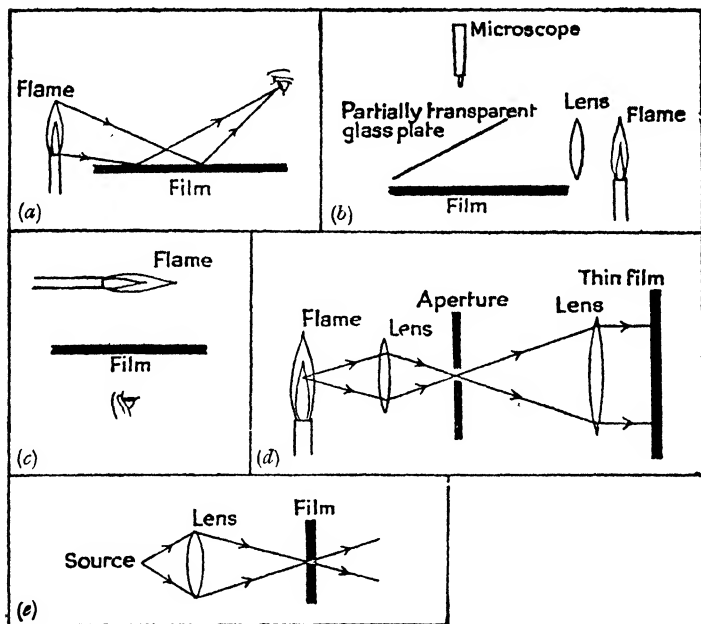
FIG. 108

In (a) and (b) the fringes are seen by reflected light. In 108 (a) the rays reaching the eye from different parts of the film come from different parts of the sodium flame and are incident

on the film with different values of cos r. In 108 (b) the rays which enter the microscope must all have hit the surface approximately at the normal, and the variation in path difference ($2\mu e \cos r$), and any fringes seen, will depend solely on the thickness e, and not on the term $\cos r$. The fringes can, therefore, be regarded as being contour lines showing the thickness of the film, each line being separated from its neighbour by a film thickness of $\left(\dfrac{\lambda}{2\mu \cos r}\right)$, or approximately $\left(\dfrac{\lambda}{2\mu}\right)$. Since these fringes are seen by reflected light they are generally unsatisfactory if the plates are silvered, since their 'visibility' (i.e. the contrast between light and dark) is very poor.

Figs. 108 (c), (d), and (e) show normal arrangements for viewing the fringes by transmitted light—which is far the most satisfactory method with silvered films. In (c) both cos r and e may, in general, vary, but in (d) the rays are all parallel, so that cos r is constant and any lines seen must be 'contour lines' of the thin film. This is generally the best arrangement to adopt.

§ 95. Circular Fringes

In Fig. 108 (e) the thickness is constant and the angle cos r alone alters. If the fringes formed in this latter example are focused on to a screen by an auxiliary lens they will form circles concentric with the axial line; for each circle cos r will be constant, and the path difference will be $2\mu e \cos r$ and must also be $n\lambda$.

Students almost invariably make the mistake of numbering the circle from the centre. Actually the circles should be numbered from the position where the path difference is zero, i.e. from $r = 90°$ and cos (r) = o. The total thickness of the film e can therefore be determined by counting the number of rings from $r = 90°$ to $r = 0°$. In general, the thickness will not be such as to produce exactly an integral number of

rings, but will involve a fraction of a ring. This fraction can readily be found because it is the ratio of the value of cos r for the last ring to the average change in cos r for each ring of the system.

The experiment is frequently performed as an exercise to determine the thickness e of the film. Fig. 109 is a diagrammatic representation of the experimental arrangement; the second lens (which may often be dispensed with if the angle of the film is very small) forms the rings on a screen. The numerical values of cos r for the different rings are shown in
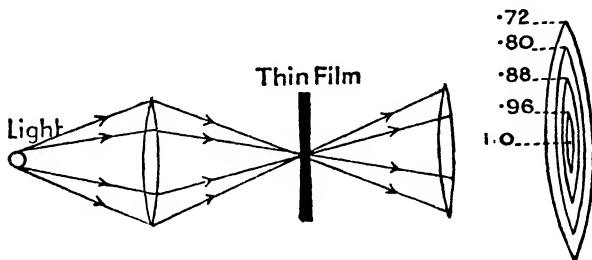
FIG. 109

the diagram, and it will be seen that the average change of cos r for each fringe is 0·08 so that the innermost ring must be $\left(\dfrac{\cdot 96}{\cdot 08}\right)$ = the 12th.

The path difference when cos $\theta = 1$ must, therefore, correspond to $12 + \dfrac{\cdot 04}{\cdot 08}$ rings, and so

$$2\mu e = (12 \cdot 5)\lambda.$$

§ 96. Newton's Fringes

A very useful method of producing a thin film of air is by placing a convex glass lens (of long focal length) on a sheet of plane glass (Fig. 110). The contour lines representing the thickness of the air film are then circles and the fringes seen

with light normal to the circle are called Newton's rings. Now if R is the radius of curvature of the curved surface,

$$x(2R - x) = d^2,$$

so approximately $$x = \frac{d^2}{2R},$$

and the total thickness of the air film will be $\left(x_0 + \frac{d^2}{R2}\right)$.

The interference fringes will be found whenever $2\mu e \cos r = n\lambda$, i.e. when

$$2\mu\left(x_0 + \frac{d^2}{2R}\right) = n\lambda \quad \text{since} \quad \cos r = 1.$$
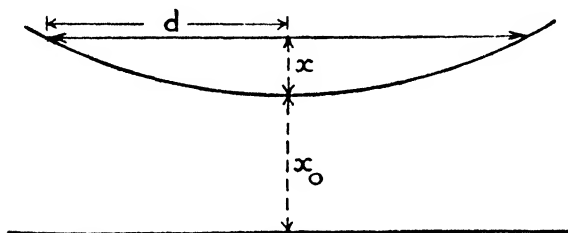
Fig. 110

The fringes are commonly employed as a means of measuring tiny displacements of the bottom plate relative to the top one. These displacements may, for example, be produced by the compression of a metal or stone pillar when it is loaded with weights and could be used to determine Young's modulus for such specimens. In a similar way the thermal expansion of small crystals relative to some standard material such as quartz has been measured.

This method of measuring small displacements is, therefore, a general one and should be remembered by the student whenever he is devising any new experiment involving such small displacements. For all these purposes it is necessary to determine not only the total change in the number of fringes or rings but also the change in the *fraction of the ring*. This can easily be done because it is the ratio of the square of the

diameter (d_0) of the smallest ring to the average change in d^2 for each ring.

§ 97. The Effect of silvering Thin Films

In all of these experiments on thin films the possible accuracy may be improved immensely by silvering the two reflecting surfaces and by viewing the films with transmitted light instead of reflected light.

The practical results obtained by silvering the surfaces of the film are that the bright parts retain their intensity, whilst
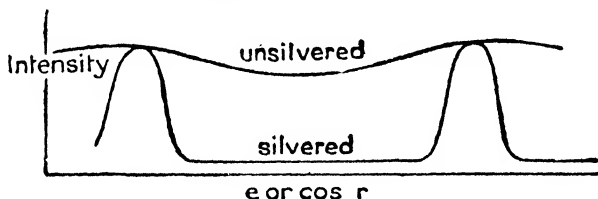
Fig. 111

the intensity of the dark parts decreases almost to zero from being nearly equal to that of the bright rings. Besides this improvement in 'visibility' of the fringes, it is found that the width of bright rings decreases and becomes a very small fraction of the width of the dark fringes. The difference between these two types is summarized by Fig. 111.

This double advantage produced by silvering the surfaces of thin films is of vital importance, and in the Fabry and Pérot interferometer, the banded spectrum experiment, and in most other practical applications of Newton's fringes, the accuracy of the measurement can be increased by a factor of at least ten by silvering the plates, and the resolving power can be increased by a factor of nearly one hundred.

The theoretical explanation is of considerable interest. The fringes seen with transmitted light are due to the combination of terms of the type $r \cos pt$, $ar \cos (pt + \delta)$, $a^2r \cos (pt + 2\delta)$, etc., as may be seen from Fig. 112. The change

in character of the fringes observed with transmitted light is due entirely to the increase in the ratio a to its limiting value of unity. This will occur also with reflected light, *except for* the special ray r_1, which, alone, will far outweigh in intensity the whole of the contributions from the remaining components, by a difference which will be increasingly greater the heavier is the deposit of silver on the surfaces.
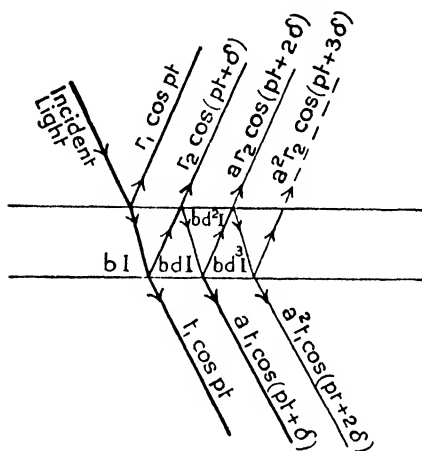
FIG. 112

The whole effect formed by reflected light will thus be identical with that seen with transmitted light with the exception that both light and dark parts of the fringe system will be overwhelmed with this great amplitude r_1. Thus, although the bright fringes do get smaller with increasing silvering, this advantage is completely offset by the very great diminution of visibility or contrast between light and dark fringes.

Returning to a consideration of the fringes produced by transmitted light: the total light amplitude at the focus of the lens employed to bring together all of the parallel rays will be represented by

$$t \cos pt + at \cos (pt+\delta) + a^2t \cos (pt+2\delta) + a^3t \cos (pt+3\delta) + \ldots$$

These components must be compounded by the vectorial method, which is later mentioned in connection with the theory of diffraction (p. 208). Two vectorial representations for unsilvered films are shown in Fig. 113 (*a*) and (*b*) and have small values for δ. The remaining two representations, (*c*) and (*d*), are for a silvered film and have the ratio a very near to unity.

These diagrams should make it clear, in a general way, that if the ratio a is small, a very considerable variation of δ from values of $2\pi n$ will be required to produce an appreciable diminution of intensity, and also the amplitude will not drop to zero for any value of δ. The diagrams also show that if the ratio a is near unity the slightest variation of δ from $2\pi n$
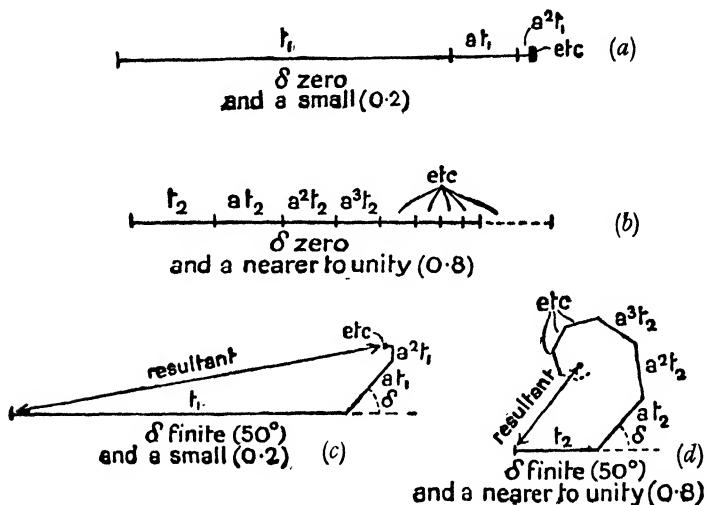
FIG. 113

will reduce the straight amplitude line of Fig. 113 (c) to a spiral (d) which passes very close indeed to the origin.

Exact mathematical calculation can be carried out which confirms this general impression. It shows that an increase in silvering makes the visibility (or contrast between the bright and dark fringes) much greater, and also makes the widths of the bright lines become smaller and smaller relative to the darker portions.

A practical illustration of this effect can easily be made by silvering two plane glass sheets in layers of increasing thickness. Thin air films constructed from such silvered sheets

will show the difference very markedly; the fringes seen with transmitted light being very indistinct where there is little silvering, whereas in the silvered regions they have very great contrast and show as narrow bright lines on a wide dark background. In this experiment it is interesting to employ a light source which emits several spectral lines; these will be clearly resolved and separated where the silvering is intense, but will be quite indistinguishable in the unsilvered parts of the film.

§ 98. The Michelson Interferometer and the Wave-Length of Light

The Michelson interferometer is a very important development from the theory of the interference of light reflected by
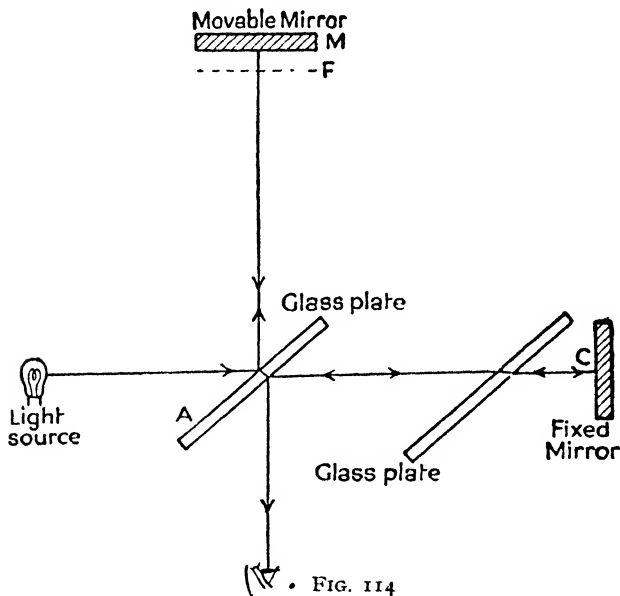
FIG. 114

thin films. Monochromatic light from a convenient source, which need not be a point, is allowed to fall on the first glass plate. *From each ray* of the incident light a part proceeds

to the movable mirror,* whence it is reflected back to the eye, and a further part proceeds to the fixed mirror and again returns to the eye as shown in Fig. 114.

The interference occurs between the part of each ray traversing the path AMA and the part traversing ACA, and the two waves will be in phase and will produce a large resultant if the optical path AMA − optical path ACA = $n\lambda$.

Correspondingly a small resultant will be produced if the optical path AMA − optical path ACA = $(n + \frac{1}{2})\lambda$.

The simplest method of understanding the action of the Michelson interferometer is to note that the eye sees the movable mirror through the glass plate, and that it also sees the fixed mirror by rays which are reflected by the glass plate. The effects seen by the eye are, there-fore, those produced by light coming from the movable mirror and from *the image of* the fixed mirror. This image is shown in Fig. 115 by the line F.
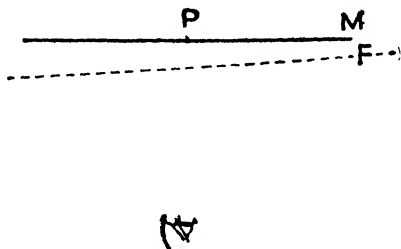
FIG. 115

The theory of the Michelson interferometer is thus resolved into the theory of the fringes seen by light reflected from two mirrors M and F, of which the one, F, is fixed, and the other, M, is movable. In Fig. 115 these two mirrors are shown at a slight angle, because it is generally impossible to arrange that the two mirrors of the interferometer are so precisely at right angles that the images will be exactly parallel with the movable mirror. Our previous discussion of the fringes produced by parallel, or nearly parallel, films, must apply here, so that, in monochromatic light, the eye will see a system of parallel light and dark fringes. These fringes will

* The second glass plate is needed in order to make the thickness of glass traversed by the two rays exactly the same. If it is not the same the colours seen with white light will be abnormal owing to the variation of the refractive index of the glass with wave-length.

occur at values of e (the thickness of the film between the two mirrors) given by

$$2\mu e \cos r = n\lambda,$$

i.e., in this example, by

$$2e = n\lambda.$$

If the thickness e of the air film is altered by a variation in the position of the movable mirror, then the fringes will move steadily across the field of view. The number of fringes passing any point P of the film will be such that

$$2(\text{change in } e) = (\text{change in } n)\lambda.$$

In general the change in n will not be integral, but will involve a fraction, but there will be no difficulty about determining
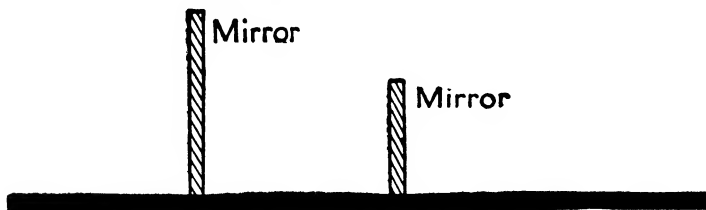
FIG. 116

this to an accuracy of about 5% or less, if the distance of the point P from the nearest bright line, and the average distance between adjacent lines, are noted.

The instrument has been used by Michelson to determine the wave-length of certain highly monochromatic lines from the cadmium spectrum. The general method of the experiment is simply to screw the movable mirror back by a known distance such as 1 cm., and to measure the number of fringes and the fraction of the fringe which crosses some particular point P on the mirrors. The precise method actually adopted is slightly more complicated on account of the great difficulty in counting the 200,000 or 300,000 fringes corresponding to a reasonable length such as 10 cm. This uncertainty as to whether the number is actually 200,000 or 200,001 is important in view of the fact that Michelson could determine the fraction of the fringe to one part in one hundred.

Michelson, therefore, constructed a number of length standards made of two mirrors (front-silvered metal ones), as shown in Fig. 116. The first was 10 cm. in length, the second one-half of this (5 cm.) as exactly as he could make it, the third was 2·5 cm., and so on, the smallest being only 0·390625 mm. in length. This smallest standard length is such that only 1,200 fringes need be counted.

Michelson placed two of these standards side by side on the carriage of his interferometer which usually supported the moving mirror. These standards are shown in Fig. 117, relative to the dotted line F which represents the position of the image of the fixed mirror.

Fringes will occur in monochromatic light between rays reflected from C or A and those reflected by F, *and it is possible*
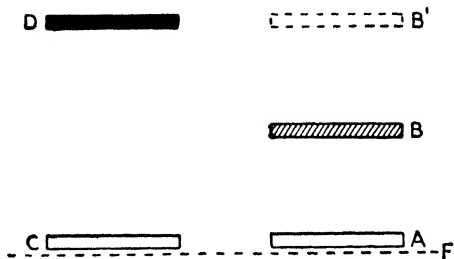
FIG. 117

to determine when either C *or* A *is exactly coincident with the image* F *of the fixed mirror because, when this does occur, we obtain a black fringe surrounded by coloured margins when the system is illuminated with white light.*

These coloured fringes are precisely the same as those obtained with white light from ordinary thin films which have already been discussed. We therefore have a means of knowing when any of the mirrors, A, B, C, or D of the standards, are exactly coincident with F.

The experiment is first conducted by illuminating with white light and adjusting both A and C to be coincident with F. The system is then illuminated with monochromatic light, and the whole of the movable carriage carrying the two standards is screwed slowly forward, the fringes crossing any point on the mirror being counted carefully until an occasional

test with white light shows that the mirror B is exactly coincident with F. The total number—and the fraction—of fringes corresponding to the small standard AB is thus determined.

The whole standard AB is then moved back until the face A is where the face B has been, and is exactly in coincidence with F. The corresponding position of B is then shown in the figure by the point B'. The whole carriage is then again screwed forward until B' is coincident with F. The total number of fringes and the fraction between A and B is therefore known and is twice that for AB; the accuracy of the determination of the fraction can naturally be increased by a direct comparison between A and B'.

The next stage is a determination of the number of fringes (and the fraction) between B' and D, so that the total fringes corresponding to CD is known, although the small portion AB alone has been counted.

The second standard, CD, of 0·78125 mm., is then compared in the same way with the third standard of 1·5625 mm., and this in turn with the fourth of 3·125 mm. Finally, we know accurately the total number of fringes in the length of the last standard (10 cm.) and can, by direct measurement, determine the fraction to the limit of accuracy.

There remains to the experimenter only the determination of the length of the last standard in terms of the international metre. This was done by Michelson with the aid of a high-magnification microscope and the laboratory standard. This laboratory standard was then compared with the national standard and this in turn with the international standard.

The final conclusion was that the number of fringes corresponding to 1 metre was, for the cadmium red, 3,106,327·0 in air at 15° C. and 760 mm. pressure. The corresponding number of light-waves in the metre was:

Red	1,553,163·5.
Green	1,966,249·7.
Blue	2,083,372·1.

There often seems to the student no valid reason why he might not repeat the experiment himself with home-made apparatus. The chief difficulty encountered in the construction of the interferometer, and the main reason for the expense of commercial types, is the extraordinary accuracy required in the screw, threads, and runners of the moving mirror, since irregularities in these should be small compared with the length 5×10^{-5} of a centimetre.

Interferometers of this type may be employed for many other purposes apart from the determination of the wave-length of light. The refractive index of gases may be determined by noting the displacement of the fringes with a variation of the gas pressure in a tube placed in one of the optical paths. The instrument may be used to measure the frequency breadth of spectral lines; or to determine the difference between the wave-lengths of spectral doublets, such as the two yellow sodium lines, by noting the thickness e of the film for which a bright fringe of the one wave-length is exactly coincident with a dark fringe of the other:

$$2e = n\lambda_1 = (n + \tfrac{1}{2})\lambda_2.$$

The student is advised to study these two questions in the usual text-books.

CHAPTER IX

DIFFRACTION

§ 99. Elementary Theory of the Diffraction Grating. § 100. The Subsidiary Maxima and Minima. § 101. Resolving Power of a Grating. § 102. More Accurate Theory of the Grating. § 103. Diffraction by a Single Slit. § 104. Diffraction by a Lens and Resolving Power of a Telescope. § 105. Diffraction by a Prism. § 106. Missing Spectra. § 107. Resolving Power of a Microscope. § 108. Propagation of Light in Straight Lines. § 109. Cornu's Spiral and Diffraction by Edges and Slits.

§ 99.. Elementary Theory of the Diffraction Grating

The spectrum produced when a diffraction grating replaces the usual prism in a spectrometer commonly provides a student with his first concrete proof that light is a wave motion with a wave-length of the order of 5×10^{-5} cm. The theory of this phenomenon can be calculated only if light consists of some form of wave radiation (the pulse theory of white light is really also a wave theory since any pulse can be analysed into component simple harmonic motions by Fourier analysis), and the agreement between the results of this theoretical calculation and those obtained by experiments may be taken as evidence in favour of the wave hypothesis.

A grating is usually assumed to consist of a glass plate on which have been ruled a large number of equidistant parallel and opaque strips or lines of equal width. In the usual arrangement, a beam of parallel light falls normally on to the grating, so that, on the other side, the transparent slits act as tiny sources which are exactly in phase with each other. In Fig. 118 the incident radiation is represented as a parallel beam of light with places on a line perpendicular to the direction of motion *all* in the same phase. On the further side of the grating we have waves emitted by each slit *in all*

directions, if the slits are of the order of 10^{-5} cm. in width. This emission of the light in approximately all directions by the slit may be taken here as an experimental fact, but is discussed in a later paragraph.

Each slit may be regarded as emitting one wave in each direction, so that, from the N slits, we must consider N waves. It is shown in the diagram that at one particular angle the
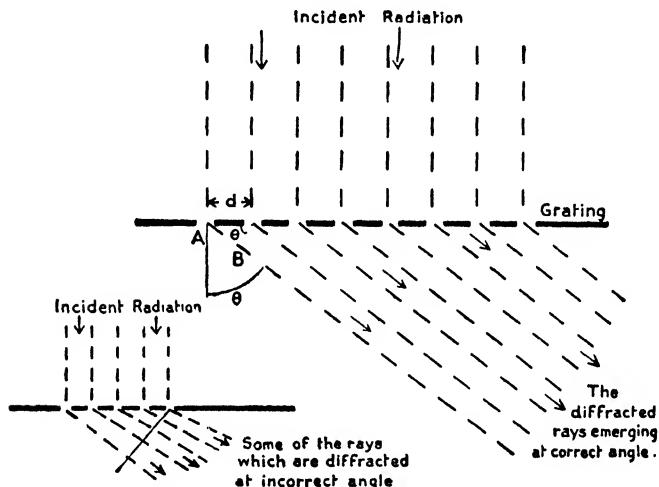
FIG. 118

phases of these N waves will be identical over any line drawn at right angles to the direction of propagation. In the usual experimental arrangements this light then falls on to a convex lens which brings all the N parallel rays to a focus without altering the relative phase differences between them.* It therefore follows that if all the N rays have the same phases over any line at right angles to the direction of propagation they will have the same phases at the focus, all their

* A lens must have this property because it brings a plane wave from an infinitely distant object to a focus, and at the focus all parts of the wave must be in the same phase or there would be much less light intensity.

amplitudes will be positive, and will all add up to give a large resultant amplitude. Thus if a bright line is produced in the focal plane of the lens, the length AB must contain *exactly* one wave-length, or, in general, exactly n wave-lengths. Thus, if d is the grating space,

$$n\lambda = d \sin \theta.$$

The usual numerical value for d is $1 \cdot 8 \times 10^{-4}$ cm., corresponding to 14,000 lines to the inch. The values of $\sin \theta$ for the violet and the red respectively are then $(0 \cdot 244)n$ and $(0 \cdot 33)n$. The angles are $14°$ and $19°$ respectively for $n = 1$, $29°$ and $41°$ for $n = 2$, and $48°$ and $90°$ for $n = 3$, while for n greater than 3, the ratio $\dfrac{n\lambda}{d}$ would be larger than unity and θ would be imaginary.

§ 100. The Subsidiary Maxima and Minima

It is easy to see in a general way that the bright images of the different spectral orders ($n = 1, 2, 3$, etc.) obtained with monochromatic light should be very sharp, narrow in shape, and should be separated from each other by a series of maxima and minima which are generally too faint to be observed.

Radiation proceeding at an angle ϕ with the normal will have a path difference of $Nd \sin \phi$ between waves from the first and last slit of the grating.* If this path difference were $Nn\lambda$, the angle ϕ would be the angle θ of the nth main spectrum. If, however, this path difference had been $Nn\lambda + \dfrac{\lambda}{2}$, for example, the amplitude of the resultant wave would have been greatly reduced because the light from the first slit would have completely interfered with that from the last (i.e. the Nth) slit, that from the second slit would have *almost* completely interfered with that from slit (N − 1),

---

* We use ϕ here in contrast to θ because θ was employed for the angle of deviation of the main spectral line, and we are here talking of rays at angles different from this, which do not form the main bright line.

and that from the third slit would have *partially* interfered with that from slit (N — 2). The total resultant amplitude would, therefore, not have been reduced completely to zero because the interference between the pairs of slits becomes less and less the nearer the pairs are together, although this total amplitude would certainly have been much smaller than when all of the slits were exactly in phase.

At a slightly different angle the path difference between the first and the last slit would be $Nn\lambda + \lambda$. Here, if N were even, the light from the centre of the grating would be exactly out of phase with that from the first slit, and the whole of the first half of the grating would interfere completely with that from the second half, so that no light at all would be emitted in this direction. If, instead, N had been odd, the whole of the first $\frac{N-1}{2}$ slits would have interfered completely with the next $\frac{N-1}{2}$ slits and the total amplitude would be that contributed by one slit alone. Since each individual amplitude is so small, however, the contribution of this slit may be ignored.

If the phase difference between the first and the last slit had been $\frac{3\pi}{2}$ the grating could have been divided into 3 equal parts. The light from A (Fig. 119) would be exactly out of phase compared with that from B and the whole of the light from

FIG. 119

AB would interfere with that from BC, so that the part CD alone would be left active.

Now this part CD is exactly analogous to the whole grating when the path difference between the ends was $\frac{\lambda}{2}$, with the exception that the active length—and therefore the amplitude —is now $\frac{1}{3}$ as great.

If the phase difference between the first and the last slit had been $Nn\lambda + 4\frac{\lambda}{2}$ the total amplitude would have been zero (or that from one or two slits only). This can again be seen by dividing the whole grating up into four equal parts which interfere with each other completely, in pairs. Similarly, if the additional path difference had been $5\frac{\lambda}{2}$, the grating would have been divided into five equal parts, of which the first four would interfere completely with each other in pairs, while the fifth alone would have contributed.
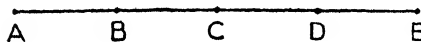
A   B   C   D   E

FIG. 120

The total amplitude will thus be one-fifth as great as when the total additional path difference was $\frac{\lambda}{2}$. (See Fig. 120).

The theory thus predicts a variation of the total resultant amplitude with the total path difference between the first and the last slit of the general character shown in Fig. 121.

$\frac{\pi}{4}a$

a

$\frac{1}{3}a$

$\frac{1}{5}a$

$+\frac{\lambda}{2}$ $+\lambda$ $+2\lambda$ $+4\lambda$

$Nn\lambda$
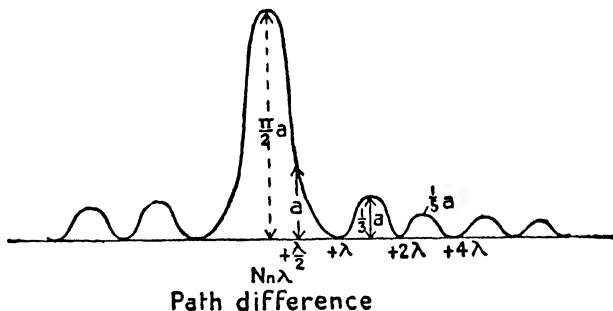
Path difference

FIG. 121

Now, in Chapter V dealing with the polarization of light (page 109), reasons are deduced for believing that the energy, or intensity, of a light wave is proportional to the square of the amplitude. The graph of energy against path difference will,

therefore, be similar to the above, but will have the large peaks still larger in comparison with the smaller ordinates.

It is next interesting to express this general graph in terms of the angle (ϕ_b) of deviation of the radiation having a path difference of $\left(n\mathrm{N}\lambda + \dfrac{b\lambda}{2}\right)$, instead of expressing it in terms of the total path difference. This can be done because the path difference $\left(\mathrm{N}n\lambda + \dfrac{b\lambda}{2}\right)$ is equal to ND sin ϕ_b, and so ϕ_b can be calculated for each value of b. The curves are nearly—but not quite—symmetrically disposed about the main bright spectral line.

We shall see later that the maximum intensity of the first peak should be only $\frac{1}{25}$ of that of the central spectral line. It is, therefore, not readily observed in a spectrometer, partly because of this low intensity, and partly because it is so close to the main maximum that it is usually hidden by the overlapping consequent on the finite width of the slit.

The magnitude of this angular separation $\delta\phi$ of the first minimum from the main peak maximum is usually calculated for the first minimum as follows:

$$\text{path difference} = \mathrm{N}d \sin \phi,$$

and so $\delta(\text{path difference}) = \mathrm{N}d \cos \phi\delta\phi.$

For the first minimum the small increment of path difference is $2\dfrac{\lambda}{2}$, so that

$$\delta\phi = \frac{\lambda}{\mathrm{N}d \cos \phi},$$

or approximately 5×10^{-5} radian, or $2\cdot5 \times 10^{-3}$ of a degree of an arc.

§ 101. Resolving Power of a Grating

If two radiations differing slightly in wave-length fall simultaneously on the grating the two main spectral lines

will be separated by a small amount. Expressed mathematically, we have

$$n\lambda = d \sin \theta,$$

so that the angular separation $\varDelta\theta$ between the two rays will be connected with the difference in the wave-lengths $\varDelta\lambda$ by the equation:

$$n\varDelta\lambda = d \cos \theta \, \varDelta\theta.$$

The graph connecting the light intensity of these two lines with the angle will then be of the character shown in Fig. 122.
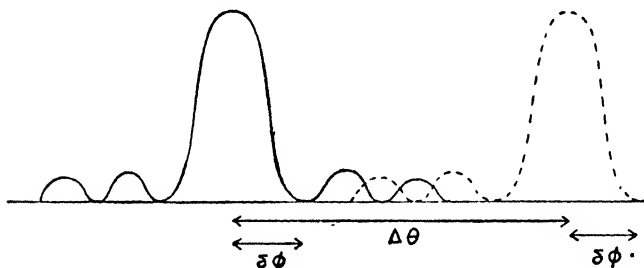
FIG. 122

It is necessary at this stage to understand quite clearly what we mean by the two small angles $\varDelta\theta$ and $\delta\phi$. The small angle $\varDelta\theta$ is given by the equation:

$$\varDelta\theta = \frac{n\varDelta\lambda}{d \cos \theta},$$

and represents the angular separation of the centres of the two bright lines produced by two radiations differing in wave-length by $\varDelta\lambda$. The small angle $\delta\phi$ is given by the equation:

$$\delta\phi = \frac{\lambda}{Nd \cos \theta},$$

and represents the angular distance between the centre of either line and its first intensity minimum.

If the difference in wave-length, $\varDelta\lambda$, between the two radiations is reduced, the value of $\varDelta\theta$ will become correspondingly smaller, and a stage will be reached at which the

two spectral peaks will fuse together so that it will no longer be possible to distinguish the separate spectral lines. This will occur in practice at different values of $\Delta\theta$ with different observers, but, in order to compare the utilities of different gratings in this repect, we agree conventionally that it will occur with a perfect observer *when the main peak maximum of the one is exactly coincident with the first minimum of the other*.

This will occur when

$$\Delta\phi = \Delta\theta.$$

It therefore follows that we cannot separate spectral lines whose wave-lengths differ by the particular value of $\Delta\lambda$ given by

$$\frac{n\Delta\lambda}{d\cos\theta} = \frac{\lambda}{Nd\cos\theta},$$

or

$$\Delta\lambda = \lambda\frac{1}{Nn}.$$

The ratio of λ to this particular value of $\Delta\lambda$ is called the 'resolving power' of the grating and is therefore (Nn). It follows that, the larger are the number of the lines in the grating and the order of the spectrum, the more readily will it be possible to distinguish two adjacent radiations.

The writer once saw a very interesting error made in this connection. In order to increase the resolving power of a spectrometer, to permit the resolution of the two sodium D lines, a very long and expensive grating with a correspondingly large number of rulings had been purchased. It was, however, used with a spectrometer whose collimator illuminated less than one-third of the total number of grating lines. The resolving power was thus governed entirely by the size of the collimator and two-thirds of the purchase price of the grating was completely wasted. The mistake was made still worse by the fact that the collimator slit was very wide and could not be adjusted, so that the size of the spectral lines was essentially that of the slit and was not dependent appreciably on the diffraction grating.

§ 102. More Accurate Theory of the Grating

In this section we obtain a mathematical expression for the exact shape of the intensity path difference curves surrounding the main spectral lines by means of a very powerful vectorial method of treatment. This vectorial method is of particular importance, since it is used for three separate purposes—the theory of a grating, Cornu's spiral, and the interference of light reflected from a thin film.

FIG. 123
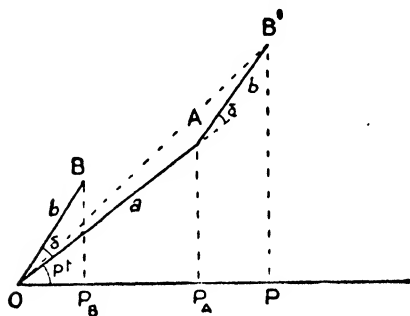
Consider two wave motions of the same frequency, but of different amplitude and phase, acting simultaneously at the same point. The total amplitude will be the arithmetical sum of the two components and may, therefore, be represented mathematically as:

$$a \cos pt + b \cos (pt + \delta).$$

Now, terms of the character ($a \cos pt$) may be represented by the distance from the centre to the foot of a perpendicular, drawn from a point which is rotating at constant speed about this centre. Thus in Fig. 123

$$OP_a = a \cos pt,$$

and

$$OP_b = b \cos (pt + \delta).$$

It therefore follows that

$$a \cos pt + b \cos (pt + \delta) = OP.$$

But

$$OP = OB' \cos (pt + \beta)$$
$$= \sqrt{a^2 + b^2 - 2ab \cos \delta} \cos (pt + \beta).$$

In addition this shows us the truth of the formal theorem which states that: 'If the amplitude of two wave components

are represented in magnitude by the lengths of two lines, and in phase difference by the angle between these two lines, then the resultant will be represented in magnitude by the length of the line joining the two ends, and in phase difference by the angle between this line and the base.
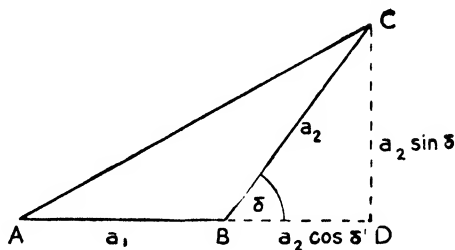
FIG. 124

The theorem, therefore, asserts that in Fig. 124 the length AC represents the magnitude, and the angle CAB represents the phase difference of the resultant.

Now $\quad AC^2 = AD^2 + CD^2$
$$= (a_1 + a_2 \cos \delta)^2 + (a_2 \sin \delta)^2,$$
which we have shown is the value of the square of the resultant.

In addition, $\tan CAB = \dfrac{CD}{AD} = \dfrac{a_2 \sin \delta}{a_1 + a_2 \cos \delta}$,

FIG. 125

which, we have shown, is the value of the tangent of the resultant phase difference.
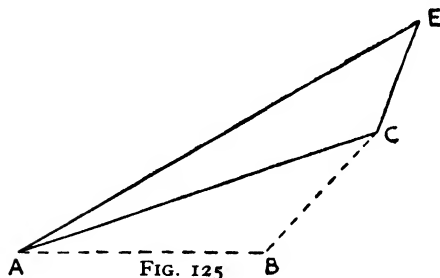
We have, therefore, proved the validity of the theorem for the combination of two simple harmonic waves. It may be extended to cover the combination of three or more waves by combining the resultant amplitude AC with that represented by EC. The final resultant of the three will then be represented in magnitude and in phase by AE (Fig. 125).

H

This extension may again be repeated by combining AE with a fourth amplitude and this resultant in turn with a fifth, so that in general the final chord will represent in magnitude and in phase the resultant amplitude of any number of waves.

Returning to a consideration of the diffraction grating, the path difference of the different waves arriving on the line
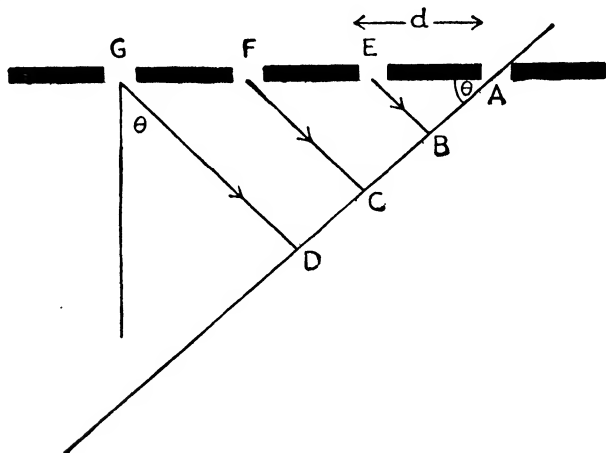
FIG. 126

ABCD (Fig. 126) are respectively o, EB, FC, and GD, i.e. o, $(d \sin \theta)$, $(2d \sin \theta)$, $(3d \sin \theta)$, etc. The corresponding *phase differences* are, therefore, o, $\frac{2\pi}{\lambda}(d \sin \theta)$, $\frac{2\pi}{\lambda}(2d \sin \theta)$, etc.*

For convenience in writing we will call these respectively o, δ, 2δ, etc.

The total resultant of all these components can then be represented mathematically by

$$a \cos pt + a \cos (pt + \delta) + a \cos (pt + 2\delta) + \text{etc.}$$

(Notice that these separate amplitudes are all equal, because

* For reasons see sections on wave motion, though the matter is of no great importance here.

we are really considering the energy from each slit sent out between the constant range θ and $\theta + d\theta$.)

If the separate components are represented graphically in the manner discussed above, they will form part of a regular polygon since the amplitudes and the phase differences are all equal (Fig. 127).
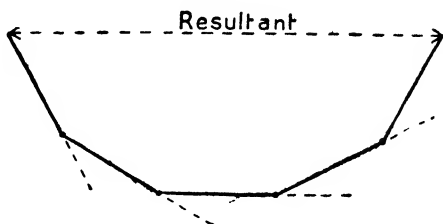
FIG. 127

Now in a real grating the separate amplitude contributions are very small, so that the regular polygon becomes part of a circle if we are considering only cases in which the phase differences are also very small. This will also occur when the phase differences are near that corresponding to one of the spectral lines, i.e. when the phase differences are $(2\pi \times n \pm$ a small quantity $\delta)$.

The final resultant amplitude is then represent⁻ by the length of the c' joining the two end⁻ portion of the circ' the phase diff' exactly zer $2\pi n)$ th⁻ straigh⁺ 128 ⸍
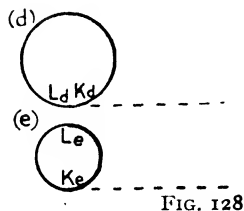
FIG. 128

phase difference the resultant in⁺ K_bL_b of Fig. 128 (b).

With a still larger phase difference the diagram will be that of Fig. 128 (c), then that of Fig. 128 (d), and then of Fig. 128 (e), and so on.

Notice that in each figure the total length of the arc remains constant because it is always the constant sum of the N separate amplitudes from each slit, and is Na.

The total angle between the tangents to the curve at K and at L respectively represent the sum (Nδ) of the phase differences between each of the N slits of the grating. It therefore represents the phase difference between the first and last slit of the grating.

Fig. 128 (c) shows that when this total phase difference between the first and last slit (Nδ) is exactly π, the amplitude is the diameter of the circle (2R), whereas when Nδ was exactly zero as in Fig. 128 (a) the amplitude was the whole line K_aL_a, which we have seen is the length of the *arc* K_bL_b or K_cL_c, etc.

The amplitude in Fig. 128 (c) is therefore

$$\frac{\text{diameter } K_cL_c}{\text{arc } K_cL_c} = \frac{2R}{\pi R}$$

of its maximum peak value.
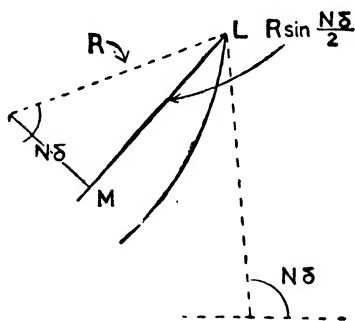
Fig. 128 (d) shows that when Nδ = 2π the total amplitude is zero, and Fig. 128 (e) shows that when Nδ = 3π the amplitude ratio is

$$\frac{\text{diameter } K_eL_e}{\text{arc } K_eL_e} = \frac{2R_e}{3\pi R_e} = \frac{2}{3\pi}$$

of its maximum value.

In Fig. 129 we consider Fig. 128 (b) in more detail. The actual amplitude received is repre-

ˌhile the total maximum amplitude

ˌe difference was the arc KL.

The ratio of the two amplitudes is

$$\frac{\text{chord KL}}{\text{arc KL}}.$$

The phase difference $N\delta$ between the first and the last slit is represented as shown, so that the chord KL is $2(\text{ML}) = 2R \sin \frac{N\delta}{2}$, while the arc KL is $R(N\delta)$.

The amplitude ratio $\dfrac{\text{chord KL}}{\text{arc KL}}$ is therefore

$$\frac{2R \sin \dfrac{N\delta}{2}}{R(N\delta)} = \frac{\sin\left(\dfrac{N\delta}{2}\right)}{\left(\dfrac{N\delta}{2}\right)}.$$

This formula for the amplitude ratio will clearly apply at the maxima and minima positions as well as at the intermediate values. When $\dfrac{N\delta}{2}$ is small, the ratio tends to unity. When $\dfrac{N\delta}{2}$ is $2\dfrac{\pi}{2}$ the sine — and the ratio — is zero. Similarly, when $\dfrac{N\delta}{2}$ is $3\left(\dfrac{\pi}{2}\right)$, the ratio is $\dfrac{1}{3\dfrac{\pi}{2}}$, when $\dfrac{N\delta}{2}$ is $4\dfrac{\pi}{2}$ it is again zero, and when $\dfrac{N\delta}{2}$ is $5\dfrac{\pi}{2}$ it has another maximum value $\dfrac{1}{5\dfrac{\pi}{2}}$.

The ratio of the intensities is proportional to the square of the amplitude, so that the intensity of the first maximum is $\left(\dfrac{2}{3\pi}\right)^2$ ($= \frac{1}{23}$ approximately) of that of the main bright line of this spectral order. Similarly, the amplitude of the second

maximum will be $\left(\dfrac{2}{5\pi}\right)^2$ ($= \frac{1}{60}$ approximately) of that of the central bright line.*

This theoretical curve of intensity against path difference has been given in Fig. 121. At considerable distances from the main peak maximum of the undeviated line ($n = 0$), or of the different bright spectral lines ($n = 1, 2, 3$, etc.), the curve obtained from the formula will not be rigidly accurate, because the individual phase differences between adjacent slits are here so large that the amplitude polygon cannot be regarded as a circle even to a first approximation.

§ 103. Diffraction by a Single Slit

The approximation involved in the above theory would be rigidly valid only if the grating had consisted of an infinite number of slits each contributing an infinitely small amplitude. The phase differences between adjacent slits would then have been so small that the whole infinite number of them would have produced the total phase difference between light coming from the two edges of the grating. Now if the individual size of the grating space were infinitely small, real spectra could not be found, because

$$n\lambda = d \sin \theta,$$

and d would be smaller than λ.

Thus there would be no possibility of distinguishing between this grating, with an infinite number of infinitely fine slits, and a piece of plane glass of the same size; in fact, the amplitude circle of the previous section—and the results obtained from it—can apply rigidly only if the grating is replaced by plane glass, so that the transparent parts of the grating are completely continuous.

* The calculations given here neglect the fact that the real maxima of the intensity graphs occur slightly before $\dfrac{N\delta}{2} = \dfrac{\pi}{2}$ because the radius of the circle is then bigger. This correction is, however, very small.

It therefore appears that if a beam of parallel light falls on to a rectangular glass window and is then brought to a focus by a lens, the central bright focal spot will be surrounded by a system of weak subsidiary rings according to Fig. 130.
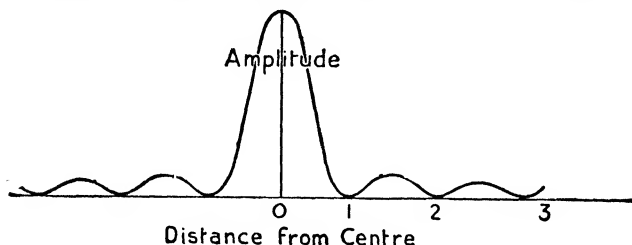
FIG. 130

The different minima of intensity will occur at all places at which $\dfrac{\sin^2 \dfrac{N\delta}{2}}{\left(\dfrac{N\delta}{2}\right)^2}$ is equal to zero. This, we have shown, occurs whenever the total phase difference between light coming from the two ends of the grating or aperture ($N\delta$) has the values 2π, 4π, etc.

There will, therefore, be zero light intensity at all those angles with the incident radiation for which the corresponding path difference is λ, 2λ, etc. Fig. 131 shows that these angles are given by the equation:

$$n\lambda = D \sin \theta,$$

where D is the total width of the aperture.

The numerical value of θ for the first intensity minimum is exceedingly small for apertures of finite size. Thus, if D = 1 cm., $\sin \theta$ is of the order of 5×10^{-5}, and θ is of the order of 2.5×10^{-3} degrees.

The smaller the numerical value of D is, the larger will be the angular deviation of the first minimum; if

$$D = \tfrac{1}{10} \text{ mm.}, \ \theta = 0.25°;$$
and if
$$D = 10^{-3} \text{ mm.}, \ \theta = 30°.$$

Since these diffraction maxima and minima can be seen readily only if the light is passing through a very narrow aperture, the pheno- menon is usually spoken of as the 'diffraction pro- duced by a single slit,' and is fre- quently calculated separately from the corresponding cal- culation for the grating. As we have seen above, however, the small diffraction maxima and minima pro- duced by a grating are merely the dif- fraction effects pro- duced by one single aperture of the same size as the whole grating.
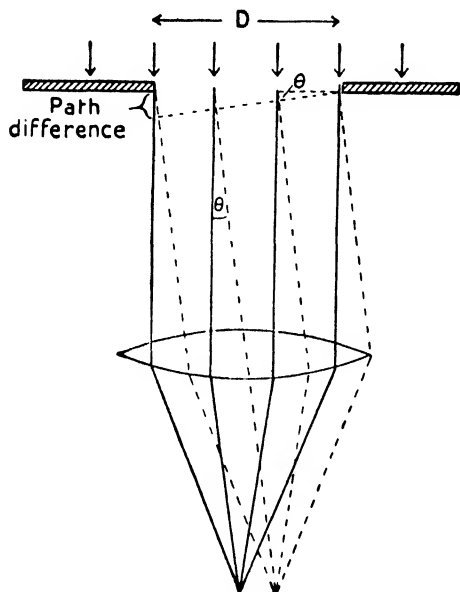
FIG. 131

§ 104. Diffraction by a Lens and Resolving Power of a Telescope

This is another example of a diffraction phenomenon which is conventionally obscured by being given an entirely separate treatment. The radius of the diffraction patch produced at the focus by perfectly parallel light entering the aperture and the lens in Fig. 131 is clearly $F\theta$, i.e. it is $F\dfrac{\lambda}{D}$ for the first intensity minimum if F is the focal length.

This means that a distant star cannot form a point image in the focal plane of a telescope objective, but must form instead a diffraction patch which will be smaller, the larger is the diameter of the aperture or of the lens. If the lens

diameter is increased not only does the size of the diffraction patch diminish, but also the energy of the light entering the lens will increase. The apparent brightness of the image of the star, therefore, increases proportionally to the cube of the lens diameter. This is the main advantage in employing telescopes with very large aperture objectives in astronomical work.

The actual numerical value of the size of the diffraction patch $F\dfrac{\lambda}{D}$ given above is necessarily not quite accurate, so that the real size is greater or less than this by a small numerical factor of the order of 1·22. The value $F\dfrac{\lambda}{D}$ (and $\sin\theta = \dfrac{n\lambda}{D}$ for a slit) really applies rigidly to a plane wave incident to a long rectangular aperture, and converged to a line focus by a cylindrical lens. If instead we use a spherical lens and converge the rays to a point, we clearly expect to obtain a slightly different numerical value. This criticism applies particularly to the telescope, because, in addition, the aperture is circular here instead of rectangular.

§ 105. Diffraction by a Prism

Here again a separate treatment is given in the usual text-books, and here again the same result for the angular width of the refracted monochromatic radiation may be obtained directly by considering the prism as a rectangular aperture producing a normal diffraction pattern. This angular width of the beam is given by $\sin\theta = \dfrac{\lambda}{D}$ if we include the whole of the first diffraction maximum out as far as the first minimum.

Now

$$\mu = \sin \tfrac{1}{2}\frac{(A + \delta)}{\sin \tfrac{1}{2}A},$$

and

$$d\mu = \frac{\cos \tfrac{1}{2}(A + \delta)}{\sin \tfrac{1}{2}A}\,(\tfrac{1}{2}d\delta),$$

*H

and once more we may assume that we shall be unable to distinguish two monochromatic radiators if the peak maximum of the one lies on top of the first minimum of the other, so that we may calculate the value of $d\mu$, for which $d\delta = \dfrac{\lambda}{D}$.

$$d\mu = \left(\frac{\cos \frac{1}{2}(A + \delta)}{\sin \frac{1}{2}A}\right)\left(\frac{\lambda}{2D}\right).$$

The resolving power, which is the name given to the ratio $\dfrac{\lambda}{\delta\lambda}$, can now be found for the prism, and is

$$\frac{\lambda}{d\lambda} = \frac{d\mu}{d\lambda}\frac{\sin \frac{1}{2}A}{\cos \frac{1}{2}(A + \delta)}(2D).$$

Normally it will suffice to leave this expression in the form given and to observe that the resolving power increases with the dispersion of the prism, and with the size of the prism. If, however, one is interested in simple geometry, the expression may be simplified by considering Fig. 132 in which the rays are proceeding *symmetrically*.
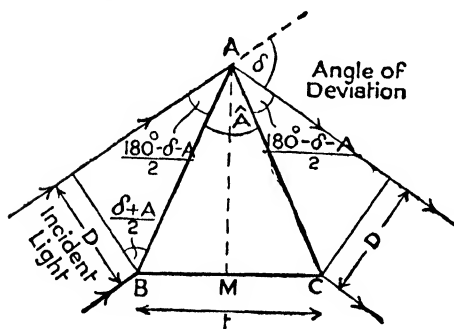
FIG. 132

Clearly

$$\frac{(2D) \sin \frac{1}{2}A}{\cos \frac{1}{2}(A + \delta)} = 2(AB) \sin \frac{1}{2}A = 2 \times (BM, = t.$$

the length of the base.

The resolving power of a prism is thus

$$\frac{\lambda}{d\lambda} = \frac{d\mu}{d\lambda} \times t.$$

§ 106. Missing Spectra

When spectra are formed by a diffraction grating, it is found that certain orders (for example, $n = 2$) are sometimes absent or are present with greatly diminished intensity. This is a direct consequence of the diffraction effects produced by each of the tiny slits that go to make up the complete grating. The matter may be made clearer by an example.

The first minimum of the diffraction pattern of a single slit occurs, as we have shown, at an angle θ given by $\sin \theta = \dfrac{\lambda}{D}$, where D is the width of the slit.

Regarding the whole grating as a large slit, we see that

$$\sin \theta = \frac{\lambda}{D} = \frac{5 \times 10^{-5}}{3} = 1 \cdot 7 \times 10^{-5},$$

and $\qquad \theta = (9 \times 10^{-4})°,$

if we take the reasonable figure of 3 cm. for the width of the grating.

The value of D for the individual slits which go to make up the complete grating may be taken as $1 \cdot 8 \times 10^{-4}$ cm., and so

$$\sin \theta = \frac{5 \times 10^{-5}}{1 \cdot 8 \times 10^{-4}} = 0 \cdot 278.$$

so that $\qquad \theta \sim 16°.$

Each slit of the grating will, therefore, send no radiation at this angle; hence, if one of the spectral orders should occur here, it will have zero intensity.

The mathematical condition that the bth, $2b$th, etc. (b is any integer), spectrum is missing may be found as follows:

FIG. 133

The smallest angle θ at which any individual slit AB (Fig. 133)
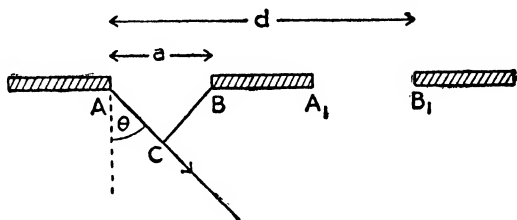
of the whole grating produces zero intensity is when the path difference between light going from A and B in the direction considered is exactly one wave-length, λ. The light from A will then interfere completely with that from M and the whole of the half AM with the whole of the half MB.

$$\therefore \ \lambda = a \sin \theta.$$

Now the bth spectrum will be missing (b is any integer we like to consider) if it is formed in this particular direction, i.e. if it has *its sin* θ also $\dfrac{\lambda}{a}$,

but
$$b\lambda = d \sin \theta,$$

$$\therefore \ b\lambda = d\frac{\lambda}{a},$$

and so the bth spectrum will be missing if the ratio $\dfrac{d}{a}$ is equal to b.

The commonest occurrence is for the grating rulings to be approximately equal in width to the space between the
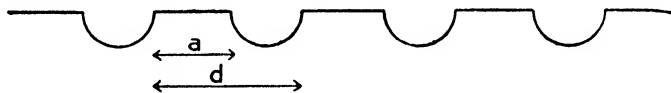
Fig. 134

rulings. The ratio $\dfrac{d}{a}$ is then two, and so the second, the fourth, the sixth, etc., spectra will be missing (Fig. 134).

This is not necessarily such a disadvantage as it at first appears, because it means that the incident light is concentrated more in the direction of the first spectrum, where it is usually required.

§ 107. The Resolving Power of a Microscope

A further interesting consequence of the diffraction effects produced by fine slits is to be found in connection with the resolving power of the microscope.

We will assume that a high-power microscope is being

employed to examine a fine slit, or a natural object which includes in itself a fine slit aperture. Now we have already shown that light from a fine slit spreads out at different angles θ and forms diffraction maxima and minima. The exact shape of the graph connecting the intensity and the angle of this diffracted light depends on the character of the slit or slit system (Fig. 135).
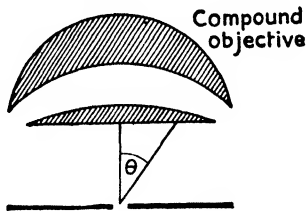
FIG. 135

Thus we are already familiar with the diagrams of Fig. 136 (a), (b), and (c). In a similar way, each different type of fine object has a different type of intensity/angle curve, which is a characteristic of this object. Thus, for example,
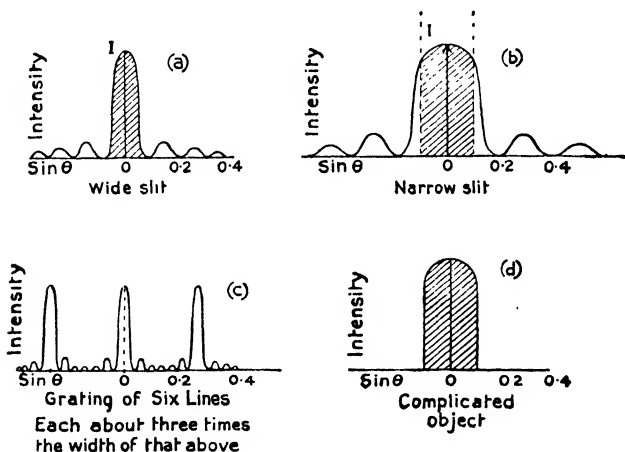
FIG. 136

Fig. 136 (d) is a diffraction picture of an object which is probably somewhat complicated.

It follows from these considerations that, when we are dealing with objects which are so small that we cannot assume the propagation of light in straight lines, the image is

produced by the microscope objective from the diffracted rays, and a different image results if the intensity of these diffracted rays varies with the angle θ in a different manner. Suppose now that the lenses of the microscope objective were only able to collect to a focus rays coming at not more than 20° with the normal. Then, in this example, the diffracted rays from the fine slit of Fig. 136 (*b*) would have the angular distribution shown by the shaded area. Now this distribution is precisely the same as that coming from the 'complicated' object of Fig. 136 (*d*). In consequence the image formed would be a reproduction of this complicated object, and would not be the simple narrow slit.

In order to compare the utilities of different microscopes from this point of view it is usual to assume that the image will be completely different from the object if the whole of the first diffraction maxima down to the first minimum does not enter the eye. Now this first minimum occurs at an angle θ given by

$$\lambda = d \sin \theta,$$

if d is the slit width.

It follows from this that the angular aperture of the objective must be as great as possible if it is to be employed to magnify objects of a size comparable with λ. Now the maximum value of $\sin \theta$ is unity, so that even if the angular aperture of the object glass has the limiting value of 90°, it will still not be possible to obtain accurate images of objects whose sizes are less than the wave-length of light.

The usual text-books give the practical arrangements— such as oil immersion and aplanatic points—by which the angle of collection of the diffracted rays may be increased.

§ 108. The Propagation of Light in Straight Lines

It is well known that one of the greatest diffculties encountered by the wave theory of light used to be in connection with the experimental fact that light travels in straight lines

in any homogeneous medium.   This seemed in marked contrast to the behaviour of sound or water waves, which were known to bend round corners to a very considerable extent.

We now know that this difference is a consequence of the large differences between the wave-lengths of light and of water or sound waves.   Experiments with ripple tanks will show, for example, that a wave will bend very readily round
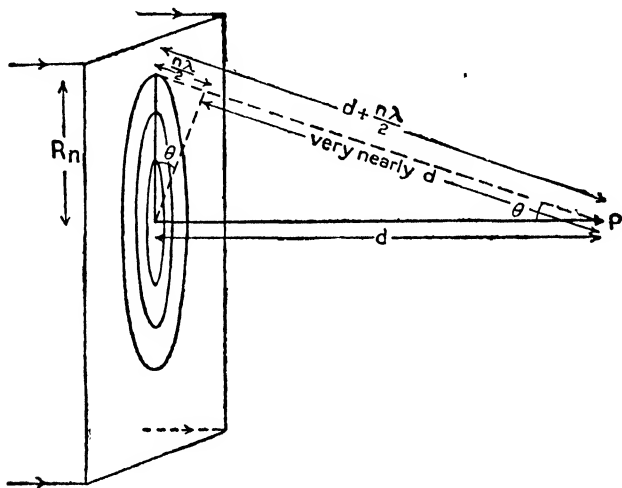
Fig. 137

a single edge if its wave-length is large, but that if the wave length is small this bending is materially reduced.

The usual treatment of this subject considers a parallel beam of light with a plane wave front moving towards a point P (Fig. 137).

A series of circles are then drawn on this plane wave front, so that the distance from the edge of the circles to the point

P increases in steps of $\dfrac{\lambda}{2}$.

The radius of the nth ring, R_n, can readily be found because, from the diagram,

$$\frac{n\frac{\lambda}{2}}{R_n} = \sin \theta,$$

which is also very nearly equal to $\frac{R_n}{d}$.

Therefore

$$R_n{}^2 = \frac{n\lambda}{2}d.$$

The usual treatment then goes on to show that the areas between successive circles are all equal, because

$$\pi R^2{}_n - \pi R^2{}_{n-1} = \pi\left(n\frac{\lambda}{2}d - (n-1)\frac{\lambda}{2}d\right) = \pi\frac{\lambda}{2}d.$$

The amplitude contributions from all of these annular rings are, therefore, all equal, or—if increasing obliquity and distance be considered—all diminish very slowly indeed. Further, since the path differences of light from different circles increases in steps of $\frac{\lambda}{2}$, the sign of these contributions from the separate annular rings will alternate. Thus the total amplitude observed will be the sum of a large number of gradually decreasing values, and

total amplitude $= a_1 - a_2 + a_3 - a_4 + a_5$, etc.

This is then usually written as

$$\frac{a_1}{2} + \left(\frac{a_1}{2} - a_2 + \frac{a_3}{2}\right) + \left(\frac{a_3}{2} - a_4 + \frac{a_5}{2}\right) + \text{etc.},$$

which is equal to $\left(\frac{a_1}{2} \pm \frac{a_n}{2}\right)$ if the summation is only continued to n terms. (The sign will be positive if n is odd and negative if n is even.) This result follows directly if the terms $\left(\frac{a_1}{2} - a_2 + \frac{a_3}{2}\right)$ are all taken to be exactly zero, which will be accurately so if the amplitude contributions A_1, A_2, etc., are decreasing proportionally to n.

It is usual, at this stage, for the student to grow tired and to close his essay with a rather vague feeling that this has

explained why light travels in straight lines. The discussion must, however, be carried materially beyond this stage, and he should consider at least one *numerical* example to show the general accuracy with which light should be expected to travel in straight lines. The simplest of these examples is that in which the beam of parallel light, with its plane wave front, meets a sharp absorbing edge and forms a shadow of this on a screen (Fig. 138).
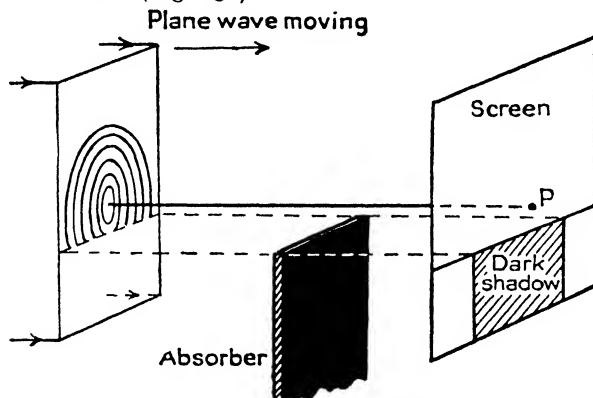
FIG. 138

Clearly, in this example, if light really does travel in straight lines we shall have an absolutely discontinuous change in intensity at the shadow boundary from complete zero to the full maximum.

Now consider the illumination at the point P which is just above the absorbing edge. The light from the first three annular rings will be reaching P without encountering the absorbing edge. Light from the remaining annular rings will encounter the absorber and will be reduced in amplitude by a greater and greater extent. Thus, the total amplitude for this point may be represented by values of the type

$$A_1 - A_2 + A_3 - 0.95A_4 + 0.90A_5 - 0.84A_6$$
$$+ 0.78A_7 - 0.72A_8 - 0.67A_9 + \ldots,$$

where the numerical values have been guessed.

It therefore follows that this intensity will be somewhat different from its value if the absorber had not been present, and that, to this extent, the light has not been propagated in straight lines.

If the point P had been slightly lower and nearer to the geometrical shadow the amplitudes could have been written:

$$a_1 - 0{\cdot}95a_2 + 0{\cdot}89a_3 - 0{\cdot}81a_4 + \text{etc.},$$

which will be appreciably different from the previous value.

It follows from these considerations that the intensities will vary until the points considered are so far distant from the geometrical shadow that a number of the annular rings or half-period elements reach the point without encountering the absorber. We cannot, at this stage, say accurately how many of these half-period elements will be needed before the light intensity becomes approximately constant, but it seems probable that this number will be more than one and less than one hundred.
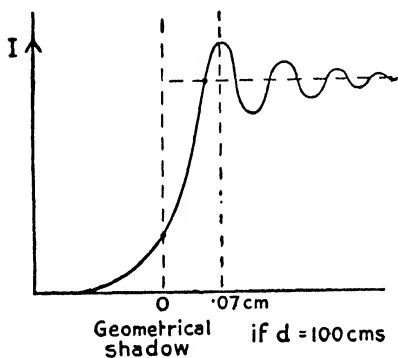
Fig. 139

Now if the distance d between the absorber (or wave front) and the screen is one metre, the radius of the tenth ring, for example, is given by

$$R_{10}^2 = \frac{10}{2} \times 100 = 10 \times 2{\cdot}5 \times 10^{-5} \times 100 \text{ cm.}$$

and $\quad R_{10} = 1{\cdot}6 \times 10^{-1} \text{ cm.}$

Thus the theory predicts that the intensity will be constant at distances of the order of 1 mm. from the geometric shadow when the absorbing edge is one metre distant.

The general vagueness of this result is so unsatisfactory that the author rarely troubles to explain this theory to his

students, but prefers to 'kill two birds with one stone' by investigating the more accurate theory of diffraction by edges and slits, by showing that the intensity/distance diagram for the diffraction of a plane wave by a straight edge considered above is of the type shown in Fig. 139, and by pointing out that light travels in approximately straight lines because the numerical scale of this diagram is so small.

The above method has, however, one very big advantage in that it gives a very clear insight into the operation of zone plates and diffraction by circular obstacles. These matters are very important and are missed here, not because they are not needed, but because they are carefully discussed in all the usual text-books.

§ 109. Cornu's Spiral and Diffraction by Edges and Slits

Consider a plane wave front moving on to the point P (Fig. 140). Divide this wave up into strips of length dx at different distances x from the normal from P.
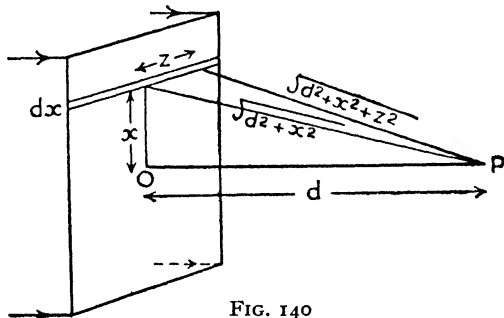
FIG. 140

The path difference of light coming from a distance z along this strip is clearly

$$\sqrt{d^2 + x^2 + z^2} - d,$$

which is

$$d\left(1 + \frac{x^2 + z^2}{d^2}\right)^{\frac{1}{2}} - d = d\left(1 + \tfrac{1}{2}\frac{x^2 + z^2}{d^2}\right) - d = \frac{x^2 + z^2}{2d}$$

to a first approximation.

It follows from this that the total effect at the point P *from each single strip* will be the resultant of all of these separate contributions from the little portions (dx, dz). Let us represent *this total resultant for the particular strip for which $x = 0$* by the equation:

$$\text{amplitude} = (adx) \cos pt.$$

Then from any other strip, at any other height x above, *the resultant* can be expressed by:

$$(adx) \cos \left(pt + \frac{2\pi}{\lambda} \frac{x^2}{2d} \right).$$

The amplitude (adx) is the same as in the first strip if it is of the same thickness dx, because we may assume that d and $\sqrt{d^2 + x^2}$ are nearly the same. There is present a phase difference $\left(\dfrac{2\pi}{\lambda} \dfrac{x^2}{2d} \right)$ corresponding to a path difference of $\dfrac{x^2}{2d}$, because *every single* point (dx, dz) of this strip has this difference in path compared with the *corresponding* point in the zero strip $(x = 0)$, so that—if every amplitude contribution has this same path difference—the resultant must also have it.

Thus the final resultant at the point P is the sum of all of the separate equal amplitudes (adx) with phase differences of $\dfrac{2\pi}{\lambda} \left(\dfrac{x^2}{2d} \right)$, which increase as we consider values of x which are progressively greater.
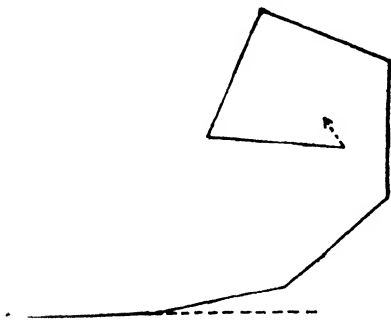
FIG. 141

Now we have considered, on page 209, a vectorial method by which we can find the total resultant amplitude of a number of components which differ in phase. If a polygon is constructed in which the numerical values of the amplitudes are

represented by the lengths of lines, and the phase differences are represented by the angles between these lines and the base, then the resultant is represented in magnitude and in phase by the length of the line joining the two ends of the polygon. In our example we are considering equal strips and therefore approximately equal amplitudes, and we are considering phase differences which increase proportionally to the *square* of the distance from the origin. The angles between adjacent amplitude lines, therefore, increase progressively in the manner shown and the curve approximates to a spiral (Fig. 141).

The smaller we take the numerical values of dx the more closely will the curve approximate to a spiral, and, in the limit, when dx tends to zero, the curve will be a perfect spiral.
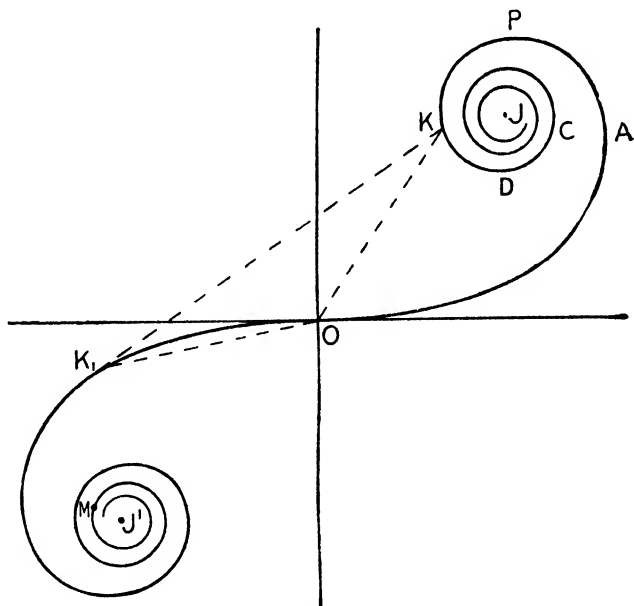
FIG. 142

In practice, if we consider values of x below the zero line, as well as above it, we obtain a double spiral—which is called Cornu's spiral. A part of this curve has been drawn accurately in Fig. 142, and we must consider its practical use.*

The Cornu spiral must be interpreted in connection with the diagram of the plane wave moving in the direction of the point P (Fig. 143). The path difference of light coming from
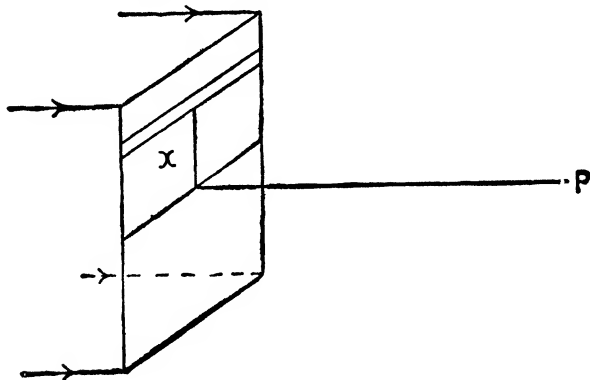
FIG. 143

a strip a height x above the base strip is, we know, $\dfrac{2\pi}{\lambda}\left(\dfrac{x^2}{2d}\right)$, so that we can calculate the value of x corresponding to any

* An approximation to this curve may be constructed arithmetically by dividing up the wave front into equal strips of lengths of the order of $\dfrac{\sqrt{\lambda d}}{50}$ or less. The amplitude lines are then all equal and the angle between the first and the second of these will be $\dfrac{2\pi}{\lambda}\left(\dfrac{\frac{\sqrt{\lambda d}}{50}}{2d}\right)$, that between the first and the third $\dfrac{2\pi}{\lambda}\left(\dfrac{\frac{2\sqrt{\lambda d}}{50}}{2d}\right)^2$, etc.

The rigid mathematical calculation turns on Fresnel's integrals and can be found in the advanced text-books, though it is not required for the purposes for which this book is written.

point, such as A, for example, on the Cornu spiral (Fig. 142). This point A corresponds to the phase difference $\dfrac{\pi}{2}$, so that

$$\frac{2\pi}{\lambda}\left(\frac{x_a{}^2}{2d}\right) = \frac{\pi}{2},$$

and

$$x_a{}^2 = \frac{\lambda d}{2}.$$

Similarly B has a phase difference of π, and so $x_b{}^2 = \lambda d$. D has a phase difference of 2π, so that the value of x_d is given by

$$x_d{}^2 = 2\lambda d.$$

It therefore follows that for any numerical value of x we may calculate the angle between the tangent and the base, and can therefore calculate the corresponding position on the Cornu spiral (Fig. 142).

Now the length OK, for example, on the Cornu spiral represents the amplitude received at the point P (Fig. 143) from all of the strips included between $x = 0$, and the value of x corresponding to the point K. Similarly, K¹ refers to light coming from below the zero line, and OK¹ represents the amplitude from this portion. In consequence the total

·07 cm if screen to edge = 100 cm
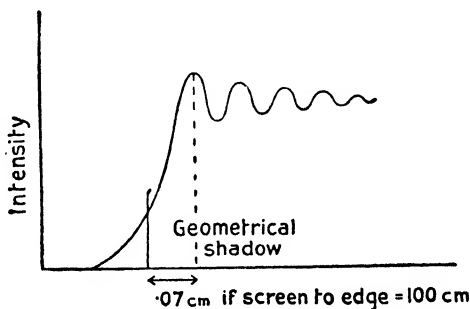
FIG. 144

amplitude from the portion below the zero line and that above it will be represented by K'K.
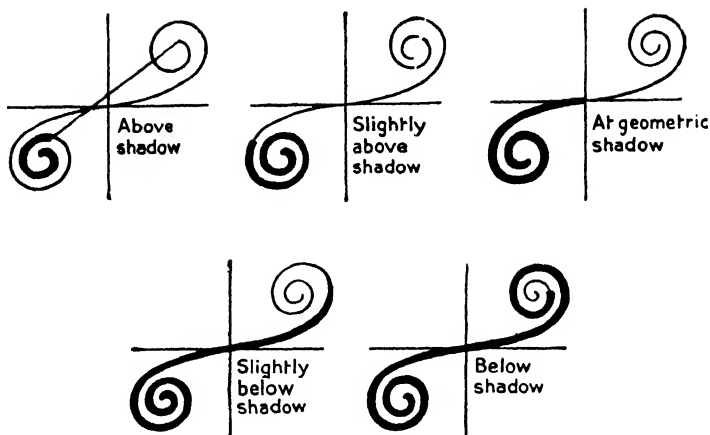
The heavy line represents the active part of the spiral

FIG. 145. CORNU'S SPIRAL FOR A PLANE WAVE
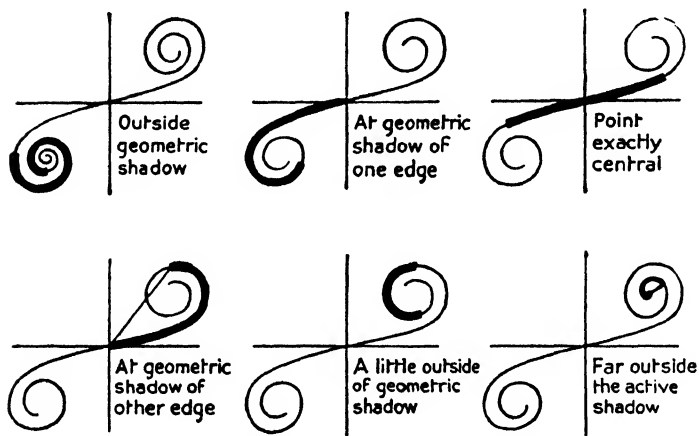
The heavy line represents the active part of the spiral

FIG. 146. CORNU'S SPIRAL FOR A SINGLE SLIT

It will now be profitable to consider a numerical calculation of the amplitude received at a point at, for example, 0·08 cm. above the geometrical shadow of a straight-edge, when the screen is, say, 90 cm. from the edge.

For this example, the phase difference $\left\{ \dfrac{2\pi}{\lambda}\left(\dfrac{x^2}{2d} \right) \right\}$ will have the limiting values ∞ and $-\dfrac{2\pi}{5 \times 10^{-5}} \times \dfrac{0\cdot08^2}{2 \times 90} = -4\cdot45 \times \pi.$

The corresponding points on the Cornu spiral can then be found—with the help of a protractor—and are J and M (Fig. 142). The total amplitude is, therefore, represented by the length of the line JM. Now if there had been no absorbing edge present, the amplitude would have been represented by JJ′, thus the amplitude ratio is $\dfrac{JM}{JJ'}$, which is 1·05. Finally, the intensity is proportional to the square of the amplitude, and so is $\left(\dfrac{JM}{JJ'} \right)^2$ of the value it could have had if the absorber had been absent.

Working in this way it is possible to draw accurately the theoretical curve for the intensity against distance from the geometric shadow. It is because of the small scale of this theoretical curve shown in Fig. 144 that we now find no difficulty in reconciling the wave theory with our experimental knowledge of the rectilinear propagation of light.
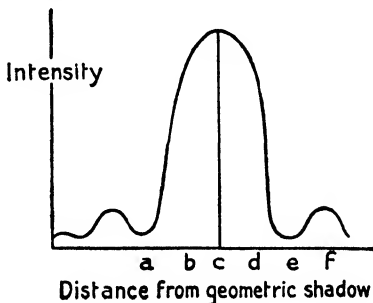
FIG. 147

The general shape of this curve can be seen immediately by considering the Cornu spirals for a series of values of P above and below the geometric shadow (Fig. 145).

A similar series of figures shows how the Cornu spiral can

explain the diffraction phenomena which are observed when a single slit is illuminated with parallel light (Fig. 146).

The intensity/distance curve derived from this latter is illustrated roughly in Fig. 147. It is naturally very similar to that of Fig. 130, but here the minima are not zero. The difference between the two curves is due to the fact that in the example of Fig. 130 a lens was placed after the slit to condense together all of the parallel rays, whereas in the present example no lens is employed, so that the bright parts will overlap into the dark minima of intensity.

Intensity
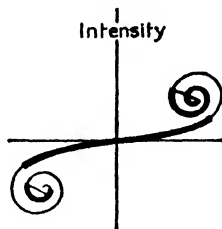
FIG. 148

The Cornu spiral can be employed to investigate theoretically the diffraction effects from a single slit or from a combination of two or more slits. These examples may safely be left as exercises for the students, once the problem has been commenced on the correct lines. This correct method may be gathered from Fig. 148, which shows the four separate amplitudes which must be combined together to represent the final amplitude received by a point which is exactly central between two absorbing lines or pins.

CHAPTER X

MICHELSON'S METHOD FOR THE DETERMINATION OF THE VELOCITY OF LIGHT

On the score of accuracy the experimental determination of the velocity of light carried out by Michelson between the years 1921 and 1926, so far surpasses any of the preceding measurements that its numerical value alone warrants consideration. The value obtained by this experiment, $c = 2 \cdot 99796 \times 10^{10}$ cm. sec., is estimated to have a probable error of only thirteen parts in a million, whereas the values obtained by other experimenters had accuracies as follows:

Fizeau	1849	Error 1 in 20	✓
Foucault	1862	Error 1 in 150	
Cornu	1872	Error 1 in 1,000	
Michelson	1879–82	Error 1·7 in 10,000	✓
Newcomb	1882	Error 1 in 10,000	

The essential principle of Michelson's method is that light is reflected from one of the mirror faces of an octagonal mirror, is transmitted to a distant reflection station placed 35 kilometres away, returns to the octagonal mirror, and is reflected into a suitable observing telescope (Fig. 149). The position of the image in the telescope is carefully noted and the octagonal mirror is made to rotate. When the speed of the mirror is sufficiently great the mirror will have turned exactly one-eighth of a revolution in the time the light takes

FIG. 149

to cover the total 70 kilometres of path. The light will then be reflected into the telescope and will form its image in exactly the same position as it was when the mirror was stationary. Hence

$$\frac{2d}{V} = \text{time for } \tfrac{1}{8} \text{ revolution} = \frac{1}{8f},$$

if f is the number of revolutions per second.

This simple scheme would be quite impracticable on account of the very small amount of energy incident on the distant

Much later

Slightly later

Ray reflected at one instant
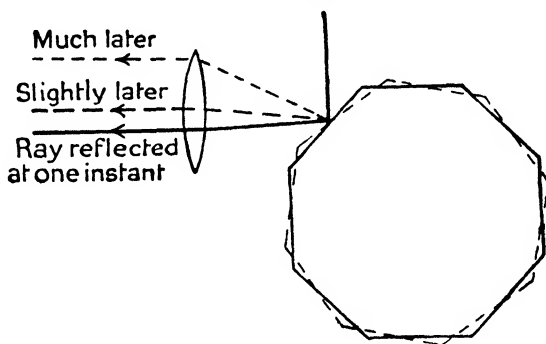
FIG. 150

reflector $\left(\text{total energy of light} \times \dfrac{\text{width of distant reflector}}{2\pi \times 35 \text{ kilometres}}\right)$ and on account of the difficulty in orientating the distant mirror to reflect back into exactly the correct position.

The first of these difficulties may be overcome by inserting a long-focus lens at a distance from the revolving mirror equal to the focal length of the lens, see Fig. 150. Thus light reflected from any particular part of the rotating mirror will always be made parallel by the lens, providing only that the angle of the mirror is such that the light does strike some part of the lens.

The second difficulty is overcome by employing any device that turns rays through exactly 180° whatever may be their angle of incidence. The cheapest of these devices is a large-

aperture concave mirror provided with a small plane (or concave) mirror at its focus. Incident parallel rays making a finite angle with the axis will be brought to a focus on the
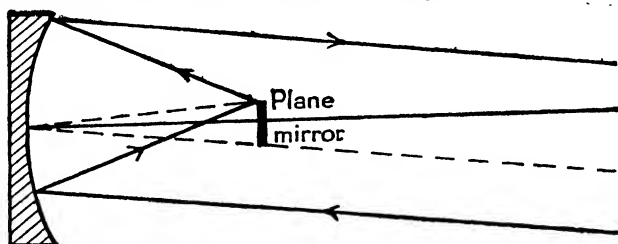
FIG. 151

plane mirror (Fig. 151), and must—after reflection—diverge from the same focus. The rays, after their second reflection from the concave mirror, must therefore all be exactly parallel to their initial direction.
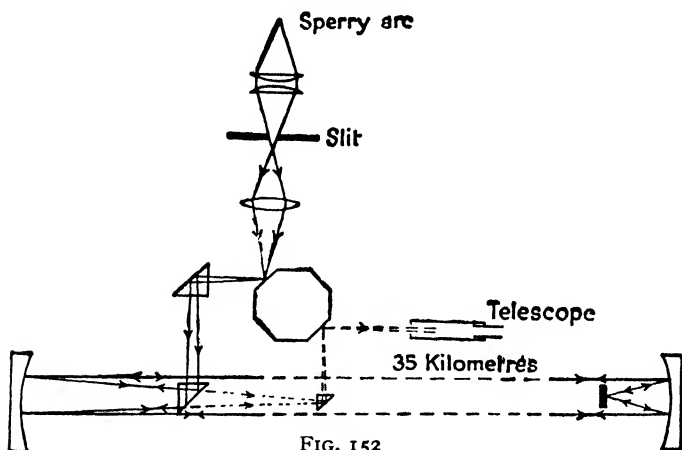
FIG. 152

In the course of his investigations Michelson varied the details of his optical arrangements. These details may be gathered from Figs. 152 and 153, from which it will be observed that the lens employed near the revolving mirror

in the above simple outline is actually replaced by a concave mirror of 30 ft. focal length.

One of the big advantages of these arrangements over the original rotating mirror of Foucault is that the second reflection is from a surface which is not bright with light scattered from the original beam. The other big advantage is that the new method is to some extent a null method: the angle of shift of the rotating mirror is known to be exactly 45°, and the accuracy of the experiment depends merely on the
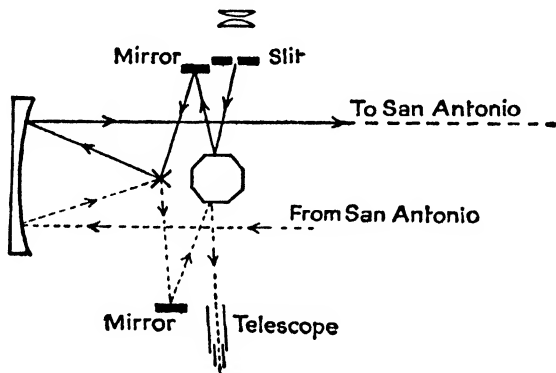
FIG. 153

accuracy with which the spot of light can be brought to its original position. Thus Michelson found it possible to determine the position of the spot to ·01 mm. on an effective radius of 250 mm. Such precision is only possible with a very small, clear, well-defined spot of light, and this in turn depends on the constancy of all the angles of the octagon. These angles were determined by an interference method and were accurate to one in a million.

The speed of the rotating mirror was 529 revolutions per second; at this speed centrifugal force is important and was sufficiently great to burst one mirror. The rotation was produced by an air-turbine attachment, compressed air from

four nozzles impinging on to a series of turbine buckets slightly below the mirror.

The speed of the rotating mirror was determined by observing the reflection of an electrically driven tuning-fork vibrating 132·25 times a second. The air delivered to the turbines was then varied until the four images of the fork seen in the mirror failed to drift across the field. At this instant the observer noted the displacement of the image, so that in the actual experiment the mirror turned

$$\{45° - (\text{a very small angle } a)\}$$

during the time the light took to travel the 70 kilometres path. The direction of rotation of the turbine was then altered so that the displacement a was on the other side of the zero line. This gave increased accuracy of measurement. The electrically driven tuning-fork was in turn compared with a free pendulum which had been carefully standardized.

The 35-kilometre optic path extended from the Mount Wilson observatory to Mount San Antonio near Pasadena. The height of these two stations prevented difficulty because of the curvature of the earth, whilst the light was travelling over a huge level plain of geologically constant character.

The distance was very carefully surveyed by the United States Survey Department, who were particularly interested on account of the possibility of using similar apparatus to determine base lines for survey work on archipelagos, or in mountainous districts where the normal 'chain' methods are unsuitable.

Corrections were made for the refraction produced by the atmosphere ($\mu = 1·000225$), an allowance being made for the variation with temperature and pressure. Finally, to reduce the irregularities in this atmospheric refraction, the experiments were normally performed one hour after sunset.

CHAPTER XI

THE GAS SCALE AND THERMODYNAMICS

§ 110. The Perfect Gas Scale of Temperature

Ordinary gases have very small intra-molecular attractions between neighbouring atoms and obey approximately the gas equation:

$$pv = R(T° C. + 273° C.).$$

The perfect gas would be one having zero intra-molecular attraction, and which followed this simple gas equation rigidly. In addition, it is generally assumed, as will be discussed later, that the specific heat at constant volume C_v and the ratio of the two specific heats, γ, of a perfect gas are constant irrespective of temperature or pressure.

Scholarship or degree candidates should be very familiar with the equation

$$pv = \tfrac{1}{3}nmV^2,$$

which is proved in all the usual text-books; n represents here the total number of molecules (each of weight m gm.) in the volume v, while V^2 represents the mean value of the squares

of the molecular or atomic velocities, and the pressure p is expressed in dynes per sq. cm.* The deduction of this equation involved the implicit assumption that there was no force of attraction between molecules, and that, during collisions, the molecules rebounded instantaneously as if they were completely concentrated into points.

It follows immediately by connecting the above two equations together that $\frac{1}{3}nmV^2$ of a perfect gas is approximately equal to R(T° C. + 273° C.) when the centigrade temperatures are determined on an ordinary 'mercury in glass scale.'

The average kinetic energy ($\frac{1}{2}mV^2$) of a perfect gas molecule, therefore, depends approximately on the 'mercury in glass temperature,' and we can *define perfect gas scales of temperature* such that the *mean kinetic energy of a perfect gas molecule is accurately proportional* to the temperature on the new scales.

We therefore have, *by definition of our new scales of temperature,*

$$pv = RT.$$

Since pv was nearly equal to R(T° C. + 273° C.) temperatures on the perfect gas scale are very nearly equal to (T° C. + 273° C.) on the mercury in glass scale, providing the size of the units on the perfect gas scale is adjusted to divide the temperature interval between the boiling-point and the freezing-point of water into 100 equal parts.

The perfect gas scale has, therefore, been constructed so that the kinetic energy of the gas is proportional to the temperature, and that

$$pv = RT.$$

It therefore follows that the ratio of two temperatures on this scale can be measured practically by the ratio of the two values of pv, by the ratio of the two volumes if the pressure is kept constant, or by the ratio of the two pressures if the volume is kept constant.

The only condition governing such determinations is that

---

* For hydrogen at N.T.P $V = 1.84 \times 10^5$ cm. per sec., $n = 2.70 \times 10^{19}$ per cc., and m—the mass of the hydrogen molecule—is 3.32×10^{-24} gm.

I

the value R must be the same for the two experiments. This in turn means that the two experiments must be conducted *with the same number of molecules*. (pv was equal to $\frac{1}{3}nm\mathrm{V}^2$, and T was proportional to $\frac{1}{3}m\mathrm{V}^2$, so that R must be proportional to n.) It is not necessary that the two experiments should be carried out with the same apparatus, or even the same gas (providing *perfect* hydrogen, *perfect* air, etc., are employed); the ratio of the temperatures will always be the ratio of the two values of (pv) providing the two experiments are conducted with the same number of molecules.

§ 111. The Specific Heat at Constant Volume

The equation $pv = \mathrm{RT}$ applies always to all perfect gases, with the numerical value of R proportional to the number of molecules. (The value of R for one molecule is $1\cdot37 \times 10^{-16}$, and the number of molecules per cc. at N.T.P. is $2\cdot70 \times 10^{19}$. The pressures are again measured in dynes per sq. cm.)

If any gas expands under conditions such that its temperature remains constant it is said to be expanding *isothermally*. During this expansion it must be doing work equal to $\int p\,dv$, which will be expressed in ergs if the pressure is in dynes per sq. cm., and the change in volume in cc. An equivalent amount of heat $\dfrac{\int p\,dv}{\mathrm{J}}$ (where J is $4\cdot2 \times 10^7$, if $\int p\,dv$ is in ergs) must, therefore, be supplied to the gas from some external source if the temperature—and therefore the internal energy —of the gas is to remain constant.

If, however, a perfect gas is heated under conditions such that it is *not* allowed to expand, *no* external work will be done, and we should expect all of the heat supplied to be spent in increasing the kinetic energy—i.e. the temperature—of the gas molecules.

We will now show that it is possible to calculate this increase in internal energy corresponding to a rise of temperature of 1° C., and thus to calculate the specific heat of the perfect gas at constant volume

Since $pv = \frac{1}{3}nmV^2$, the total kinetic energy ($\frac{1}{2}nmV^2$) of the molecules must be $\frac{3}{2}(pv)$ (which is, of course, $\frac{3}{2}RT$). If we increase the gas temperature from $T°$ to $(T + 1)°$, on the perfect gas scale, we increase the kinetic energy by $\frac{3}{2}R$, which is, however,

$$\frac{3}{2}\left(\frac{1}{T}\right)\text{(the value of } pv \text{ at } T°).$$

This change in kinetic energy will be in ergs if p is expressed in dynes per sq. cm., and if v is measured in cc.

As an illustration, we note that the volume of 1 gm. of hydrogen at N.T.P. is $11\cdot13 \times 10^3$ cc. and the pressure is $1\cdot013 \times 10^6$ dynes per sq. cm. The total increase in the kinetic energy of hydrogen when the temperature is increased from 273° to 274° should therefore be

$$\frac{3}{2}\left(\frac{1}{273}\right) \times (1\cdot013 \times 10^6 \times 11\cdot13 \times 10^3) = 6\cdot2 \times 10^7 \text{ ergs.}$$

The specific heat of perfect hydrogen at constant pressure is the number of calories required to produce this increase in kinetic energy of the 1 gm. of hydrogen, and is therefore

$$\frac{6\cdot2 \times 10^7}{4\cdot2 \times 10^7} = 1\cdot47 \text{ calories.}$$

Now, although this calculation works out correctly for monatomic gases, the actual specific heat at constant volume observed with hydrogen is $2\cdot40$ calories per gm. per °C. We may conclude from this that, with hydrogen, nearly two-fifths of the total heat supplied must go in some process other than increasing the total translational kinetic energy of the molecules.

We sometimes think that this extra energy may have been spent in warming up the individual molecules and in increasing their internal temperature as distinct from that which produced the increase in kinetic energy. This is clearly an incorrect mental picture, because we cannot talk of the temperature of a molecule separately from its kinetic energy; temperature was defined by—and is ultimately measured by —the mean kinetic energy of molecules.

A more correct explanation is that a molecule possesses other degrees of freedom from the three (x, y, z) of translational motion, and that, according to a theorem in classical statistical mechanics, called the 'Principle of Equipartition of Energy,' every single one of these extra degrees of freedom of the system should have associated with it the same energy as was involved in each of the degrees of freedom of the translational motion.

Now a diatomic molecule might be expected to vibrate in and out—that is, with the distance between the two atoms becoming alternately larger and smaller (Fig. 154). There should
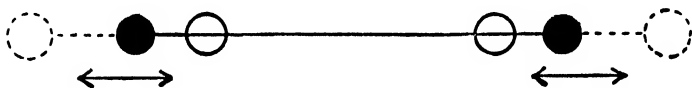
Fig. 154

be associated with this vibration two degrees of freedom—the one of kinetic and the other of potential energy (you will not probably be really convinced about this, but it is true nevertheless). Thus the total specific heat at constant volume of hydrogen should be $\dfrac{1 \cdot 47}{3}$ calories per gm. for each of the three degrees of translational freedom plus $\dfrac{1 \cdot 47}{3}$ calories per gm. for each of the two degrees of vibrational freedom. The total of $\left(5 \times \dfrac{1 \cdot 47}{3}\right) = 2 \cdot 44$ calories per gm. per °C. is sufficiently near the experimental value $2 \cdot 40$ to warrant the belief that it is the true explanation of the discrepancy.

Further reflection shows that this cannot be the complete explanation because such a diatomic molecule should possess in addition three more degrees of freedom due to its possibility of wobble in any three planes, at right angles to each other, and should possess one more degree of freedom in virtue of its possibility of spinning about the line joining the two centres (Fig. 155).

If all of these 9 degrees of freedom had been in operation the total specific heat at constant volume would have been $9 \times \dfrac{1 \cdot 47}{3} = 4 \cdot 41$ calories per gm. per °C., whereas it was found experimentally to be only $2 \cdot 4$, corresponding to a total of only five degrees of freedom. This partial disagreement with theory must mean that our classical theorem of the equipartition of energy is partially incorrect. This in turn
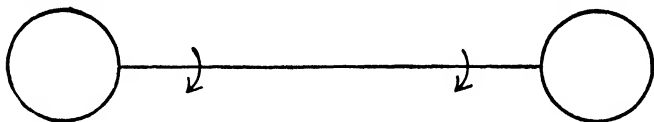

FIG. 155

seems very surprising because for many years the best theoretical physicists in the world could detect no flaw in the logic underlying it, until the quantum theory was proposed. This quantum theory effectively states that ordinary classical, Newtonian, mechanics is only approximately correct, and that, *with oscillating systems, the angular momenta* $(I\omega)$ *can only assume definite values which are always an integral number* (n) *times a fundamental constant of angular momentum, usually written as* $\dfrac{h}{2\pi}$ *(where* $\dfrac{h}{2\pi}$ *has the universal value of* $1 \cdot 042 \times 10^{-27}$ *erg sec.).*

Thus

$$I\omega = \frac{nh}{2\pi},$$

where n is an integer.

The energy of a rotating system is therefore

$$\tfrac{1}{2}I\omega^2 = \tfrac{1}{2}I\left(\frac{nh}{2\pi I}\right)^2 = \tfrac{1}{2}\left(\frac{h}{2\pi}\right)^2 \frac{n^2}{I}.$$

It therefore follows that the energy of such vibrating or rotating systems is not continuous but must vary in jumps corresponding to unit changes of n. If the moment of inertia of the system I is very great these energy changes will be so

small that they can be regarded as being continuously variable. Correspondingly, if the moment of inertia is very small the energy changes will be very large.

In the example on the specific heat of hydrogen which we were discussing above, the angular momenta of the vibrations and, in particular, of the spins are so small that the smallest energy values which could be associated with them are very great compared with the total energy of the molecule at ordinary temperatures. In consequence *very few molecules indeed* have any energy at all in these types of motion and the *change* in the number of the vibrating molecules produced by any ordinary rise of temperature is so insignificant that we may agree that neither the spin nor the vibrations are operating at ordinary room temperatures. The total specific heat at constant volume of hydrogen, therefore, corresponds only to the kinetic energy and to part of the energy of the wobbling or rotational motion and the theoretical and experimental values agree with each other.

We may summarize this paragraph by noting that

$$pv = \text{RT} = \tfrac{1}{3}nm\text{V}^2,$$

and that the total kinetic energy of the molecules is

$$\tfrac{1}{2}nm\text{V}^2 = \tfrac{3}{2}\text{RT}.$$

The energy needed to increase the kinetic energy of a mass M gms. by an amount corresponding to $1°$ is then $\tfrac{3}{2}$R, so that the heat required in calories for each gram should be $\dfrac{3}{2}\dfrac{\text{R}}{\text{MJ}}$, R is here determined by

$$(p \text{ in dynes}) \ (v \text{ in cc.}) = \text{R}(\text{T in C. gas scale units}),$$

and J is here the number of ergs per calorie $= 4\cdot2 \times 10^7$.

This heat of $\dfrac{3}{2}\dfrac{\text{R}}{\text{MJ}}$ for 1 gm. for $1°$ rise of temperature is spread over the three degrees of translational freedom. If the total number of degrees of freedom operative under the quantum theory conditions is $3 + n$, the total heat required per gm. per degree rise in temperature will be

$$C_v = \frac{\text{R}}{2\text{MJ}} (3 + n).$$

§ 112. The Difference between the Specific Heat at Constant Pressure and that at Constant Volume

The heat required to raise the temperature of unit mass of a gas through a rise in temperature dT must be greater if the gas is expanding and is doing external work at the same time as it is being heated. Clearly, the total heat supplied must be that required to increase the temperature of the molecules under conditions such that no external work is done ($MC_v d$T), plus the heat $\dfrac{pdv}{J}$, required to perform the external work which has been performed during the expansion. Thus, if the gas is heated under conditions such that its pressure remains constant, the total heat absorbed —which can be written as $MC_p d$T—can be expressed by the equation:

$$MC_p d\text{T} = MC_v d\text{T} + \frac{pdv}{J}.$$

Although, in general, $pdv + vdp = Rd$T, yet, in this particular example, we are trying to discover an expression for the specific heat of a gas *at constant pressure*. Thus for this particular example the pressure is constant and $dp = 0$,

so
$$pdv = Rd\text{T}.$$

Thus
$$MC_p d\text{T} = MC_v d\text{T} + \frac{Rd\text{T}}{J},$$

and
$$C_p = C_v + \frac{R}{MJ}.^*$$

This equation is sometimes written obscurely without M or J, and the student is warned to avoid such misleading proofs or statements. J is frequently neglected because—somewhere at the beginning of the chapter—the author mentions that he proposes to express C_p, and C_v, and all heats absorbed or rejected, in terms of ergs rather than in terms of calories. Similarly, the term M is neglected because it is also stated that the symbol R is being employed in this particular paragraph

---

* R is always found from the equation $pv = R$T, and here p is in dynes per sq. cm., v in cc., and T in degrees on the gas scale.

for the gas constant *of unit mass*. The objection to this simplified statement is that, in rapid revisions, it is difficult to keep in mind these two qualifications, so that—although the book is actually correct—the student will usually be incorrect.

§ 113. The Equation for the Adiabatic Expansion of a Perfect Gas

In the example given above we were considering changes of temperature produced by the application of heat under conditions such that the pressure remained constant, so that $dp = 0$. We will now consider another example in which the pressure, the volume, and the temperature all vary in such a way that the heat needed to perform the external work is taken from the gas itself, which simultaneously cools. Expansions or compressions of this character, in which no heat is supplied to the gas from external sources, are very common in nature whenever volume changes occur too quickly for heat to be conducted appreciably to the gas. They are called adiabatic changes.

Thus, for an adiabatic change, the external work done, pdv ergs, is produced by a corresponding diminution of the internal, or thermal, energy of the gas, $MC_v dT$ calories. Thus

$$pdv = -MC_v dT \times J.$$

$$\therefore \frac{RT}{v} dv = -MC_v dT \times J;$$

but

$$\frac{R}{MJ} = (C_p - C_v),$$

$$\therefore (C_p - C_v)T\frac{dv}{v} = -C_v dT,$$

and

$$(\gamma - 1)\frac{dv}{v} = -\frac{dT}{T}, \text{ since } \frac{C_p}{C_v} = \gamma.$$

$$\therefore \log \left(\frac{v_1}{v_2}\right)^{\gamma-1} = -\log \left(\frac{T_1}{T_2}\right),$$

and

$$\left(\frac{v_1}{v_2}\right)^{\gamma-1} = \frac{T_2}{T_1}.$$

Moreover, since $$pv = RT,$$

$$\left(\frac{v_1}{v_2}\right)^{\gamma-1} = \frac{p_2 v_2}{p_1 v_1},$$

so that $$p_1 v_1^{\gamma} = p_2 v_2^{\gamma} = K.$$

§ 114. The Clement and Désormes Experiment

The value of γ can naturally be determined experimentally by measuring the specific heat at constant volume with a Joly steam calorimeter and by measuring the specific heat at constant pressure by Regnault's method or by continuous flow arrangements.

The ratio of these two specific heats of a gas can also be determined by a separate simple experiment first performed by Clement and Désormes. Although this experiment is given in many of the usual text-books it seems important enough to repeat here.* The experiment is conducted in the following stages:

(1). A large container is pumped up with the gas considered and is allowed to remain until it is in thermal equilibrium with its surroundings. The pressure of the gas is noted by a manometer. Then

$$p_1 v_1 = RT_1.$$

(2) The valve is then opened quickly, the gas expands to a new volume v_2, and the pressure drops to that of the atmosphere. Then

$$p_2 v_2 = RT_2.$$

This expansion is governed by the fact that it is so rapid that there can be no transference of heat from the surroundings to the gas. Thus

$$p_1 v_1^{\gamma} = p_2 v_2^{\gamma}.$$

---

* A very simple and short treatment of the theory of this experiment is given in Roberts's *Heat and Thermodynamics*. The method given here seems, however, to have value as an exercise, although it is more lengthy.

*I

(3) The valve is then quickly shut, leaving some of the gas molecules outside. Because there are now fewer molecules, a smaller numerical value must be employed for the gas constant in the general gas equation. Thus,

$$p_2 v_1 = r T_2.$$

Notice that we are here using the second pressure and the second temperature because these have not had time to alter appreciably before the valve closed. The volume v_1 is employed because it is, of course, the constant volume of the container. The symbol r in the equation replaces the previous term R, because some of the gas has been lost and the gas constant must be smaller.

(4) The gas is then allowed to warm up to the temperature of the surroundings T_1, the volume still remains that of the container, and so the pressure must increase to some new value p_3. Thus

$$p_3 v_1 = r T_1.$$

This completes the 'physics of the experiment,' and there remains merely the algebraic manipulation of the above five equations to obtain γ in terms of quantities (usually the pressures) which can be measured.

Eliminating in turn R, r, $\dfrac{T_1}{T_2}$, and $\dfrac{v_1}{v_2}$, we find that

$$\left(\frac{p_1}{p_2}\right)^{\gamma} \left(\frac{p_2}{p_1}\right) = \left(\frac{p_3}{p_2}\right)^{\gamma},$$

from which we may prove that

$$\gamma = \frac{\log \dfrac{p_2}{p_1}}{\log \dfrac{p_3}{p_1}}$$

Frequently in our experiments we employ this formula, but occasionally we find it difficult because seven-figure logarithms must be used if we are to obtain γ to the accuracy that the pressure measurements should permit. This is due to the fact that in many experiments p_2 is the pressure of the

atmosphere and the pressures p_1 and p_3 are only very slightly different from this. Thus p_1 may correspond to 1,110 cm. of paraffin and p_2 to only 1,100 cm.

This use of seven-figure logarithm tables may be avoided if the logarithms (which may *normally* both be expressed to any base desired, such as each one to the base 10) are considered *for this purpose* to the base e, and are expanded * as follows:

$$\gamma = \frac{\log_e\left(1 + \frac{(p_2 - p_1)}{p_1}\right)}{\log_e\left(1 + \frac{(p_3 - p_1)}{p_1}\right)} = \frac{\dfrac{p_2 - p_1}{p_1}}{\dfrac{p_3 - p_1}{p_1}}$$

$$\gamma = \frac{p_2 - p_1}{p_3 - p_1}.$$

In this formula $(p_2 - p_1)$ is *the difference* between the pressure of the atmosphere and the original pressure in the container, while $(p_3 - p_1)$ is the difference between the original and final pressures. The actual value of the pressure of the atmosphere does not, therefore, enter into the calculation and the *ratio* of these two differences in pressures only is involved. There is no need to convert the corresponding lengths of the manometer liquid into pressure units, and γ will be merely the ratio between the two differences in lengths, i.e.

$$\gamma = \frac{l_1}{l_3 - l_1}.$$

In this experiment it is important to have the air free from moisture because the gas cools so much during its rapid

---

* Maclaurin's Theorem states that:

$$f(x) = f(0) + xf'(0) + \frac{x^2}{\underline{|2}}f''(0) +, \text{ etc.}$$

$$\therefore \log_e(1 + x) = \log_e(1 + 0) + x\frac{1}{(1+0)} + \frac{x^2}{2}\frac{-1}{(1+0)^2} + \frac{x^3}{\underline{|3}}\frac{(-1)(-2)}{(1+0)^3}$$

$$= 0 + x + \frac{x^2}{2} + \frac{x^3}{3}$$

If you are unfamiliar with Maclaurin's theorem and therefore doubt this proof, find some table of logarithms to the base e and work out one or two examples numerically. If you do this, however, make x small compared with 1.

expansion that the water vapour would to a great extent condense and form mists*; the corresponding latent heat of evaporation of this water would then be given to the gas so that the expansion would be very far from adiabatic.

This necessity for dry air is less important if p_1 is chosen to be less than the atmospheric pressure, because then the air rushes into the container and warms the gas, instead of cooling it. There is, therefore, no tendency for the vapour to condense and the compression can still be adiabatic, although the value of γ will then—naturally—be that for a mixture of air and water vapour and will not be the value for dry air.

The experiment can be conducted with gases other than air, providing precautions are taken to prevent the entrance of air into the container when the valve is opened, although—for educational purposes—it is usual to keep the experiment simple by operating only with air.

This experiment suffers from the very grave defect that the air oscillates as it rushes from the container when the valve is opened. If the valve is shut before these oscillations have ceased the pressure may be greater or may be less than atmospheric, whereas if it is shut after the oscillations have died away the temperature of the gas will have increased materially by conduction from the walls. The student may make a big improvement in his experimental work by plotting a graph between the observed value of γ and the time of opening of the valve, $\frac{1}{2}$, 1, 1$\frac{1}{2}$, 2 secs., etc. In research work the temperatures and not the pressures are recorded.

The values of γ for gases other than air are generally determined by comparing the wave-length of sound of a constant pitch in the gas with its wave-length in air. Clearly

$$\frac{\lambda \text{ gas}}{\lambda \text{ air}} = \frac{V \text{ gas}}{V \text{ air}} = \sqrt{\frac{\gamma \text{ gas}}{\gamma \text{ air}} \times \frac{\text{mol. wt. air}}{\text{mol. wt. gas}}},$$

since it is proved in the usual work on sound that the velocity

$$V = \sqrt{\gamma \frac{p}{\rho}}.$$

* On large expansions.

These experiments usually find only the ratio of the velocity in a gas to that in air, and do not determine velocities in absolute units; they are therefore relative and not absolute experiments, and find only the ratio of γ for the gas to γ for air. To determine the value of γ for our standard gas—air—we must, therefore, either use the Clement and Désormes experiment or must determine the two specific heats directly.

§ 115. Experimental Values of γ

The experimental results for γ are precisely those we should expect from our knowledge that γ is merely the name given to the ratio of the two specific heats of a gas $\dfrac{C_p}{C_v}$ and that

$$C_p - C_v = \frac{R}{MJ},$$

and that

$$C_v = \frac{R(3+n)}{2MJ}.$$

Thus

$$\gamma = \frac{C_p}{C_v} = \frac{C_v + \dfrac{R}{MJ}}{C_v} = \frac{\dfrac{R}{2MJ}(3+n) + \dfrac{R}{MJ}}{\dfrac{R}{MJ}(3+n)}$$

$$\therefore \gamma = \frac{5+n}{3+n},$$

where n is the number of degrees of freedom other than the three of translation which are operating under the quantum theory conditions.

With monatomic gases n must be zero, and so $\gamma = \dfrac{5}{3}$.

With diatomic gases n may be zero if no other degrees of freedom are operative, may be 3 if the system has a wobbling or rotational motion, may be 5 if it is also vibrating, and may be 6 if it is also spinning about its axial line. At ordinary temperatures the first of these only is usually operative and n is less than 3.

Thus, for diatomic gases,

$$\gamma = \frac{5 + 2}{3 + 2} = \frac{7}{5} = 1 \cdot 4.$$

§ 116. The Specific Heat of a Saturated Vapour

The specific heat of a gas is the heat which is required to increase the internal energy of 1 gm. of the gas through 1 C.°, plus that required to perform the external work which has been done in the corresponding expansion. Since we can let a gas expand by any amount we desire, by adjusting the pressure, this external work—and the specific heat—might have almost any value between certain limits. The most important specific heats of gases are those in which the heating is performed either with the pressure or with the volume constant. It is, however, interesting to consider *vapours* being heated under conditions such that their pressure is increased according to the vapour pressure curve. This specific heat is spoken of as the *specific heat of saturated vapour*.

When the vapour is heated it will normally expand at constant pressure, and we must compress it if we desire to keep it saturated. Heat is generated by this compression and may be either greater or less than was required to heat the vapour, so that the specific heat of saturated vapour may be negative or positive respectively.

If the pressure of saturated vapour is increased suddenly the temperature will rise by an amount given by the curve of pressure against temperature for adiabatic expansions. If this adiabatic pressure /temperature curve had happened to be identical with the vapour pressure curve, the rise in temperature produced by the sudden increase in pressure would have corresponded exactly with the temperature of saturated vapour at that pressure. No heat would then be required to change the temperature of the gas to make it saturated, and the specific heat of saturated vapour would be zero.

Similarly, if the rise of temperature produced by a sudden adiabatic increase in pressure were greater than the corre-

sponding amount read off from the vapour pressure curve for this same increase in pressure, then the vapour would have to be cooled to make it saturated, and the specific heat of the saturated vapour would be negative.

§ 117. The Intersection of Isothermals and Adiabatics

The perfect gas scale of temperature is defined so that the general gas equation $pv = RT$ is obeyed accurately, with R proportional to the number of molecules present. Since adiabatic expansions, in which no heat is absorbed or is emitted by the gas, are governed by equations of the type $pv^\gamma = K$, perfect gases should have constant values for γ whatever is the temperature or pressure. This in turn means that a perfect gas must have a constant value for C_v, since

$$\gamma = \frac{C_p}{C_v} = \frac{C_v + \dfrac{R}{MJ}}{C_v}.$$

Thus, for example, 'perfect argon' should obey the equations $pv^{1.66}=K$ and 'perfect oxygen' equations of the type $pv^{1.40}=K$.

Any isothermal ($pv = RT_1$) intersects the steeper adiabatic curve ($pv^\gamma = K_1$) at a point (a in Fig. 156) which is common to each of these equations.
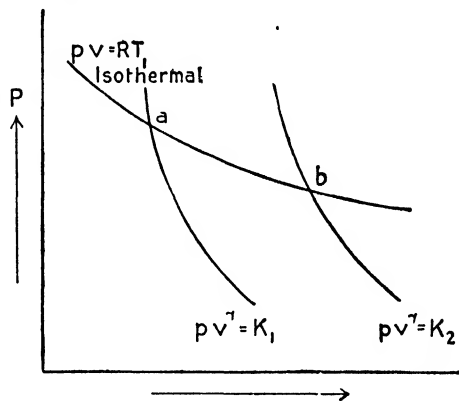
FIG. 156

The volume corresponding to this point can be obtained by removing p from these two equations, and is

$$v_a = \left(\frac{K_1}{RT_1}\right)^{\frac{1}{\gamma-1}}.$$

Similarly, the volume corresponding to the point b is

$$v_b = \left(\frac{K_2}{RT_1}\right)^{\frac{1}{\gamma-1}}.$$

(From this there emerges the interesting property of isothermals and adiabatics that the ratio

$$\frac{v_b}{v_a} = \left(\frac{K_2}{K_1}\right)^{\frac{1}{\gamma-1}},$$

being independent of T, must be the same for all isothermals crossing the two adiabatics.)
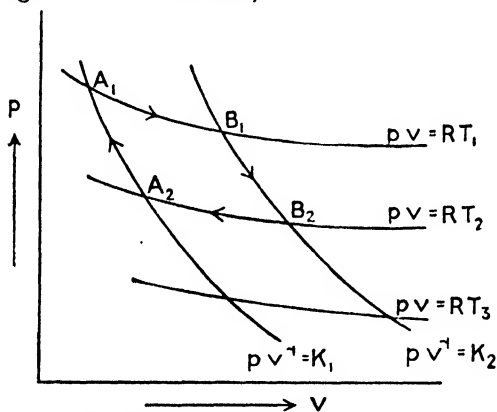
Heat absorbed or work done from A to B is proportional to Temperature on perfect gas scale

FIG. 157

We will now consider the external work done as the gas expands isothermally from the point a to the point b. This work will be expressed in ergs, if p is in dynes per sq. cm., and if v is in cc., and is clearly

$$\int_a^b p\,dv = \int_a^b \frac{RT_1 dv}{v} = RT_1 \log_e \frac{v_b}{v_a}.$$

This external work is therefore

$$RT_1 \log_e \left(\frac{K_2}{K_1}\right)^{\frac{1}{\gamma-1}}.$$

This is a very important result because it means that if a whole series of isothermals are drawn cutting the two adiabatics K_1 and K_2, then *the work done or the heat absorbed as we proceed along the isothermal from A_1 to B_1 (Fig. 157), or from A_2 to B_2, etc., will always be proportional to the temperature of the isothermal,* since the quantity

$$R \log_e \left(\frac{K_2}{K_1}\right)^{\frac{1}{\gamma-1}}$$

is constant for this quantity of gas.

Since the temperature and internal energy of a perfect gas expanding isothermally is constant it follows that *the heat absorbed by any constant quantity of a perfect gas expanding isothermally between two adiabatics is proportional to the temperature of the isothermal.*

§ 118. Thermodynamics. The Carnot Cycle

A vitally important example of this property of a perfect gas is afforded by a Carnot cycle. Starting from a point
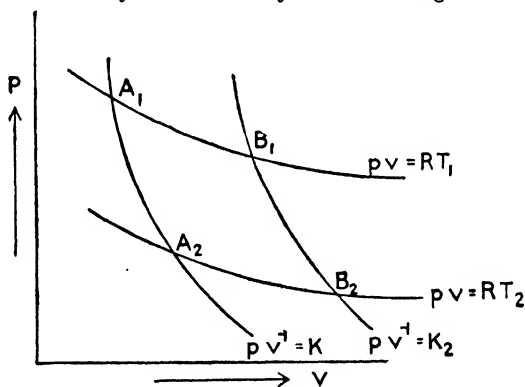
FIG. 158

such as A_1 the perfect gas is allowed to expand isothermally to B_1 and is then expanded adiabatically from B_1 to B_2. It is next compressed isothermally to the point A_2 and then adiabatically again to A_1 (Fig. 158).

Since this is a purely theoretical discussion we need not concern ourselves with the practical methods by which we could expand or compress the gas along these isothermals or adiabatics. It is, however, interesting to note that the experiment could be performed by taking a simple cylinder and piston without valves such as could be made from a bicycle pump. The cylinder should be polished externally and should have an infinitely small water equivalent.

First place the pump in a bath at a temperature of T_1 and allow the gas to expand so slowly that it absorbs heat and is kept at a constant temperature, from point A_1 to B_1. Then remove the pump from the bath and allow it to expand *rapidly* till its volume has changed to that corresponding to B_2. If this expansion is rapid enough no heat would be given to the pump from the room, the expansion would be *adiabatic*, and the temperature would, in consequence, fall to some value T_2. Then place the pump immediately in another bath which is kept at this temperature T_2. Compress the gas slowly—and therefore isothermally—to the point A_2, and complete the cycle by removing the pump from the bath again and compressing it rapidly (i.e. adiabatically) again to the point A_1.

In this complete cycle the only heat interchanges occur along the two isothermals because adiabatics are defined as being curves along which no heat is absorbed or rejected.

The heat absorbed by a gas as it expands isothermally (temperature T_1) between two adiabatics K_1 and K_2 is

$$RT_1 \log_e \frac{v_b}{v_a} = RT_1 \log_e \left(\frac{K_2}{K_1}\right)^{\overline{\gamma - 1}} \text{ ergs,}$$

as was shown on page 256.

Similarly, as the gas is compressed from B_2 to A_2 along the isothermal T_2, heat will be liberated equal in amount to

$$RT_2 \log_e \left(\frac{K_2}{K_1}\right)^{\frac{1}{\gamma - 1}}.$$

Thus, in this cycle, heat is absorbed at a higher temperature T_1, and a portion only is given out again at the lower tempera-

ture T_2. The difference between these two quantities must represent the net external work that has been done. (This statement that the external work done is equal to the heat which has been absorbed in the cycle and has not been rejected is frequently spoken of as the first law of thermodynamics.)

It is customary at this stage to consider the ratio

$$\frac{work\ done\ in\ ergs}{heat\ absorbed\ in\ ergs},$$

and to call it the *efficiency*. On this definition of the word 'efficiency' it follows that

$$\text{efficiency} = \frac{\text{heat absorbed} - \text{heat rejected}}{\text{heat absorbed}}$$

$$= \frac{RT_1 \log_e \left(\frac{K_2}{K_1}\right)^{\frac{1}{\gamma-1}} - RT_2 \log_e \left(\frac{K_2}{K_1}\right)^{\frac{1}{\gamma-1}}}{RT_1 \log_e \left(\frac{K_2}{K_1}\right)^{\frac{1}{\gamma-1}}}$$

Thus efficiency $= \dfrac{Q_1 - Q_2}{Q_1} = \dfrac{T_1 - T_2}{T_1}$.

The vital importance of this result is due to the fact that we can prove that *all perfectly reversible heat engines operating between the same two temperatures must have the same efficiency, so that the efficiency of* ALL *such reversible engines can be expressed as*

$$\text{efficiency} = \frac{Q_1 - Q_2}{Q_1} = \frac{T_1 - T_2}{T_1}.$$

§ 119. All Reversible Engines working between the Same Temperatures have the Same Efficiency

Before we can logically prove this theorem we must consider the meaning of the term 'reversible' as applied to thermodynamical cycles or to heat engines.

The working substance of a thermodynamical perfectly reversible engine must be taken round a complete closed

cycle, so that at the end of the cycle the working substance is in exactly the same state as it was at the beginning of the cycle. If this were not done we could never be certain that the external work done was really equal to $Q_1 - Q_2$, because there would always be the possibility that the internal energy of the substance had changed.

As an example of this, consider a wire being stretched by forces which are so great that the elastic limit is exceeded. When the forces are removed the wire follows the line AB and finishes with a permanent extension OB (Fig. 159). These changes can then never be made into a perfect closed cycle because we can never make the extension *and* the force together zero again. In this example the work,
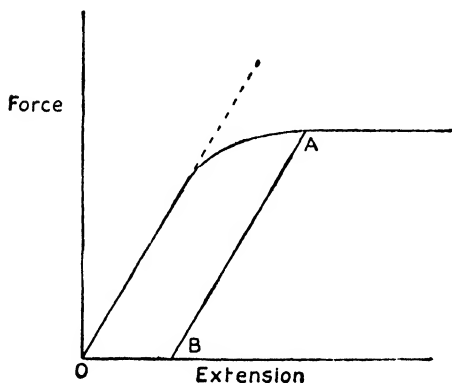
FIG. 159

which is equal to the area OAB, has increased the internal energy of the wire without the temperature having changed during any part of the cycle.

A heat engine which takes in heat at a high temperature, does external work, and rejects less heat at a lower temperature, will be reversible if it can be operated with the same efficiency in the opposite direction, taking in heat at a low temperature, giving up more heat at the higher temperature, and using mechanical energy in the process.

Although such an engine is reversible it will not be perfect if work is done at any stage of the cycle against friction, in forming eddies, etc.

In the thermodynamical sense a perfectly reversible engine is one in which every change involved in the cycle could be

reversed by an infinitely small change in the value of the forces controlling it.

It is now easy to prove that all reversible engines working between the same two temperatures have the same efficiencies by considering two perfectly reversible engines coupled together so that one takes in heat at the high temperature, produces mechanical work, and liberates less heat at the low temperature, while the other uses mechanical work, takes in heat at the low temperature, and gives out more at the higher temperature.

According to the second law of thermodynamics: *It is impossible to produce a machine which will continuously furnish useful external work by abstracting heat from the coolest body of its surroundings.*[*]

Thus, if we adjust the size of the two machines so that no heat is taken from the body at higher temperature, and $Q_1' = Q_1''$, then the second law states that no external work can be done continuously by abstracting heat from T_2. Thus, if the second law of thermodynamics is true, Q_2' for the one must equal Q_2'' for the other, so that the efficiencies

$$e = \frac{Q_1 - Q_2}{Q_1}$$

must also be equal.

This therefore shows that, if the second law of thermodynamics is true, then all perfectly reversible heat engines operating between the same temperature limits must have the same efficiencies.

Notice that this proof depends essentially on the two machines using the same high and the same low temperatures, and will not be valid if these temperature ranges differ.

§ 120. The Absolute Scale of Temperature

Scales of temperature depending on the properties of ordinary materials are arbitrary and disagree with each

[*] An alternative and equivalent statement is that 'It is not possible for any self-acting machine working in a cyclical process, unaided by any external agency, to make heat pass from one body to another at a higher temperature.'

other by amounts which become really serious at high temperatures. In this respect perfect gas scales of temperature defined by the equation

$$pv = RT$$

are much more satisfactory, because they are the same for all gases.

We have also shown that the perfect gas scales of temperature are important because temperatures are then proportional to the *kinetic* energy of the molecules. Further, providing γ, the ratio of the two specific heats of a gas, is constant, the *total* energy of the molecules of any particular perfect gas will be proportional to the temperature on this scale.

Moreover, we ·have found that the work done—or the heat absorbed—as a perfect gas is changed isothermally from one adiabatic K_1 to another K_2, is proportional to the temperature of the isothermal, on this perfect gas scale of temperature.

We have also shown that, if a perfect gas is taken round a complete Carnot cycle,

$$\text{efficiency} = \frac{Q_1 - Q_2}{Q_1} = \frac{T_1 - T_2}{T_1},$$

so that

$$\frac{Q_2}{Q_1} = \frac{T_2}{T_1}$$

and the heat absorbed or rejected on the two isothermals of a complete Carnot cycle is proportional to the temperature measured on the perfect gas scale.

This result is particularly important because all reversible engines working between the same two temperature limits have the same efficiency. Thus, whatever may be the character of the working substance, providing only that it is taken round a reversible cycle, the heats absorbed or rejected on the two isothermals are proportional to the perfect gas temperatures.

For these reasons perfect gas scales of temperature are generally spoken of as being absolute or work scales of temperature. They are independent of the properties of

any particular working substance and depend only on the properties of perfect reversible heat engines.

Notice that we have so far spoken of absolute scales of temperature, and never of 'the absolute scale of temperature'; this is because in our work with perfect gases or with heat engines we have really only needed (pv), or the heat absorbed, to be *proportional* to the temperature, and we could, therefore, have made the size of our units any value that we felt would be convenient. As an illustration, consider a perfect heat engine operating between the temperatures of condensing steam and melting ice (both at a pressure of 76 cm. of mercury). All that is known from the thermodynamics is that

$$\frac{Q_1}{Q_2} = \frac{T_1}{T_2},$$

so that the same heat ratio would have been obtained on the absolute Fahrenheit scale of temperature ($T_1 = 671\cdot4°$ and $T_2 = 491\cdot4°$) as on the absolute Centigrade scale of temperature ($T_1 = 373°$ and $T_2 = 273°$).

§ 121. The Relation between the External Pressure and the Boiling Point of a Liquid

The thermodynamical theorem that

$$\text{efficiency} = \frac{\text{work done}}{\text{heat absorbed}} = \frac{Q_1 - Q_2}{Q_1} = \frac{T_1 - T_2}{T_1}$$

finds very great practical application in a very large number of examples in which we can imagine the possibility of perfectly reversible cycles being carried out.

Thus, we may take n grams of a liquid round a complete cycle WXYZ as shown in Fig. 160.

If we keep the temperature constant and gradually increase the volume from the point W, where the whole of the substance is a liquid, vaporization will occur without any change of pressure, until all of the liquid has vaporized at the point X. In this stage, heat $= Lm$ calories has been absorbed (L is the latent heat).

We now decrease the temperature by the very slight amount dT, still keeping the vapour saturated. Since this fall of temperature is so very small we may, in the limit, neglect the heat given up in this part of the process, i.e. along XY. The substance is then completely a gas at a pressure which is slightly lower (by dp), and is represented in the diagram by the point Y.

The volume is then reduced steadily, the vapour condenses,
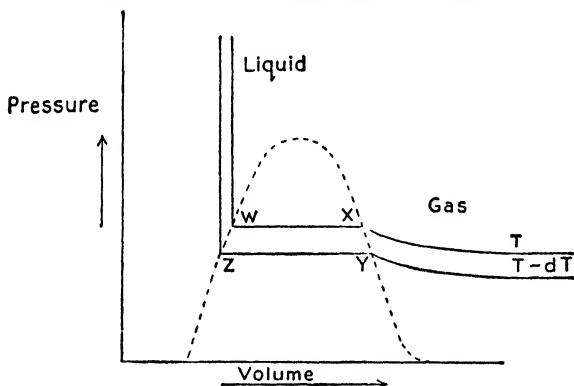
FIG. 160

and follows the path YZ till at the point Z it is completely liquid.

Finally the liquid is heated by the exceedingly small quantity of heat needed to warm it through $dT°$ back to the point W.

Now this cycle must be perfectly reversible, and so we can say that

$$\frac{\text{work done in ergs}}{\text{heat absorbed in ergs}} = \frac{dT}{T}.$$

Now the work done is very nearly the change in volume from W to X (i.e. the volume of 1 gm. of vapour — the volume of 1 gm. of liquid) times the change in pressure dp, providing we measure dp in dynes per sq. cm.

The heat absorbed is L calories (or LJ ergs) if we neglect

the tiny quantity of heat absorbed at the end ZW. Thus very nearly

$$\frac{(\text{vol. of 1 gm. of vapour} - \text{vol. of 1 gm. of liquid})\, dp}{LJ} = \frac{dT}{T}.$$

Notice that, in obtaining this formula, the approximations have been made only after we have shown that the cycle is reversible and that

$$\frac{\text{work done in ergs}}{\text{heat absorbed in ergs}} = \frac{dT}{T}.$$

Moreover, these approximations are such that they must get more and more valid the smaller we make the small change in vapour pressure dp and the small change in temperature dT, so that we are really finding the slope of the vapour pressure curve at the temperature considered.

When water changes to steam all the quantities are positive, so that $\frac{dp}{dT}$ is positive and the vapour pressure of water will increase with an increase in the temperature. The numerical value of the increase can be obtained from the formula by inserting the following figures:

T = 373° A.

L = 540 calories per gm.

Volume change from vapour to liquid = 1,676 cc.

J = 4·2 × 10⁷ ergs per calorie.

$$\therefore \frac{dp}{dT} = \frac{540 \times 4\cdot2 \times 10^7}{373 \times 1,676} \text{ dynes per sq. cm. per degree}$$

$$= \frac{540 \times 4\cdot2 \times 10^7}{373 \times 1,676 \times 13\cdot6 \times 981}$$

$$= 2\cdot72 \text{ cm. of mercury per degree,}$$

which is in agreement with the best experimental value of 2·716 cm. of mercury per degree at 100° C.

§ 122. Liquefaction

Precisely the same argument and the same formula can be employed to describe the change in melting-point with change in external pressure, and

$$\frac{dp}{dT} = \frac{LJ}{T \text{ (vol. of 1 gm. of liquid—vol. of 1 gm. of solid)}}.$$

When ice melts, the volume of the liquid formed is less than that of the original ice, and so $\frac{dp}{dT}$ is negative and the freezing-point falls as the external pressure is increased.

Here again a quantitative agreement with the experimental results is obtained.

§ 123. The Energy of a Voltaic Cell

If a voltaic cell is kept at a constant temperature, and is delivering current, external electrical work is done equal to Eq ergs, if E is the electromotive force in absolute units and if q is the quantity of electricity (in absolute units) which has been transferred.

This energy, Eq ergs, must have come from the chemical action which has occurred in the cell, and from any heating or cooling which may have taken place when the current was transferred.

We shall use the symbol H to represent the number of ergs liberated in the chemical action which is produced when unit quantity of electricity is transferred. Similarly h represents the heat which is absorbed by the cell per unit electrical transfer. Then

electrical work done = chemical energy + heat absorbed,
and $\qquad\qquad Eq = Hq + hq.$

Now suppose that we have two enclosures, the one at a temperature T and the other at a temperature $T - dT$. Let the E.M.F. of the cell when placed in these two enclosures be written respectively as E and as $E - dE$.

When the cell is in the first enclosure at T allow a transfer

of q units of electricity.   Then the heat absorbed $= qh$ ergs and the electrical work done $= Eq$ ergs.

Now put the cell in an enclosure at $T - dT$ and drive the same amount of current through it.

Then the work done *on* the cell $= (E - dE)q$.

The cycle just outlined is a perfectly reversible one and so:

$$\text{efficiency} = \frac{\text{net work done}}{\text{heat absorbed}} = \frac{dT}{T}.$$

$$\therefore \frac{dT}{T} = \frac{(dE)q}{hq},$$

and so
$$\frac{dE}{dT} = \frac{h}{T}.$$

This permits us to calculate the temperature coefficient of a cell $\left(\dfrac{dE}{dT}\right)$ in terms of the heat absorbed when the cell is transferring current, or vice versa.

Moreover, we have seen above that

$$E = H + h.$$

So
$$E = H + T\frac{dE}{dT}.$$

§ 124. Surface Tension

The surface tension of a film should properly be defined as the force acting on unit length of an imaginary cut on the film.

From this definition we may easily prove that the surface tension in dynes per cm. length is equal to the mechanical work done (in ergs) in forming unit area.*   The proof of this theorem is generally carried out by considering a $\sqcap$-shaped wire as in Fig. 161, bridged by the portion AB of length l.   If we

* These two quantities have the same dimensions, since $\dfrac{\text{force}}{\text{length}}$ is dimensionally the same as $\dfrac{\text{force} \times \text{length}}{\text{length}^2}$.

use the symbol S to represent the value of the surface tension in dynes per cm., then the force on this portion AB is lS dynes if we are considering only one surface (it is $2l$S if we are considering a film rather than a single surface). The mechanical work done as we move AB perpendicularly through a length x is (lSx) ergs. Thus the mechanical work done in forming unit area of fresh surface is
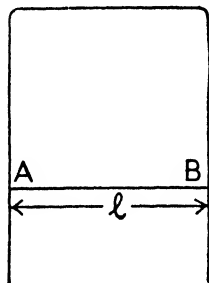
$$\frac{l\text{S}x}{lx} = \text{S ergs.}$$

It is often wrongly stated that the surface tension is the surface energy of unit area of a film surface. This is incorrect because when a film is formed heat is absorbed, so that the total surface energy is the heat absorbed plus the mechanical work done. With water the mechanical work done in forming unit area is 80 ergs, while the energy absorbed directly as heat is about 40 ergs. The theory of this heat absorption is easily discussed on the basis of thermodynamics.

Suppose h ergs is the heat absorbed when unit area of a surface is formed isothermally.

Let the surface tension, in dynes, be represented by S at the temperature T and by S $-$ dS at the temperature T $-$ dT.

If, at the temperature T, we increase the area of the film by an amount A, then the work done *on* the film will be (SA) and the heat absorbed will be (hA).

Now cool the film by an infinitesimal amount dT and then reduce the area of the film again. The work done *by* the film will be (S $-$ dS)A.

Finally heat the film again to its initial temperature T.

Then the cycle is perfectly reversible and

$$\text{efficiency} = \frac{\text{work done by film}}{Q_1} = \frac{d\text{T}}{\text{T}}.$$

Thus, if we may neglect both the infinitesimal work done

and the heat interchanges which occur while the film is being heated or cooled,

$$\frac{(S - dS)A - SA}{hA} = \frac{dT}{T},$$

and so

$$h = -\frac{TdS}{dT}.$$

Although the approximation we have made cannot be exactly true it will certainly be more valid the smaller are the values of dS and of dT, especially if the area A is kept finite. Thus, the above conclusion will be accurately correct in the limit when dS and dT are each zero.

Since the surface tension decreases as the temperature rises, it follows that $\frac{dS}{dT}$ is negative, and heat will be absorbed when a film is produced.

The total surface energy *of unit area* of a film is therefore

$$S - \frac{TdS}{dT},$$

and is numerically greater than S alone, since $\frac{dS}{dT}$ is negative.

§ 125. The Cooling of a Stretched Wire

A further interesting example of the application of thermodynamics is afforded by the stretching of an elastic wire, such as occurs in an experiment on Young's modulus.

Consider the reversible cycle ABCD (Fig. 162). The wire is first stretched by the force F and follows the path AB of the figure. The temperature is then reduced by the infinitesimal amount dT, and the wire therefore contracts by an amount dl to the point C. At this lower temperature $T - dT$ the force is gradually removed so that the wire follows the path C to D. Finally the temperature is increased to $T°$ again and the wire returns to its former state.

In this reversible cycle the mechanical work done by the

system is, in ergs, the area ABCD, which is the area AB′C′D. The mechanical work done is therefore (Fdl) ergs.

Let h be the heat absorbed in ergs at the higher temperature
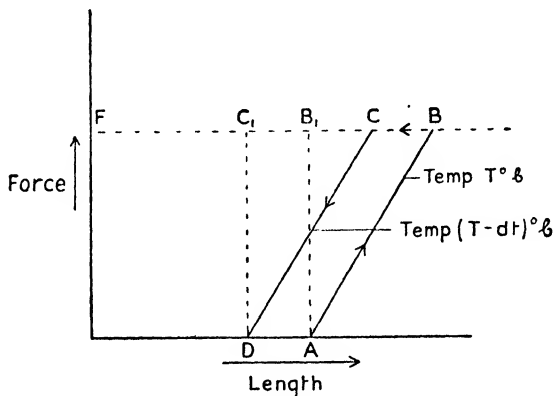
FIG. 162

when the wire was being stretched. Then, since the cycle is reversible:

$$\text{efficiency} = \frac{\text{work done } by \text{ system in ergs}}{\text{heat absorbed in ergs}} = \frac{d\text{T}}{\text{T}}.$$

$$\therefore \frac{(\text{F}dl)}{h} = \frac{d\text{T}}{\text{T}},$$

and h in ergs $= \text{FT}\dfrac{dl}{d\text{T}}.$

$$= \text{FT}(l\alpha),$$

if α is the coefficient of linear expansion and l is the normal length of the specimen. The heat absorbed in calories is then

$$h \text{ (in calories)} = \frac{\text{FT}}{\text{J}}(l\alpha).$$

Now this heat is usually taken from the wire itself and produces a cooling which may be represented by $\theta°$. Then

$$h = \theta \text{ (water equivalent of wire)}.$$

Thus, finally, we have that the wire will fall in temperature when stretched by a force F, by an amount

$$\theta = \frac{\text{FT}la}{\text{J (water equivalent of wire)}}.$$

Under normal conditions each of the quantities involved on the right-hand side of this equation are positive, so that a sudden pull on a wire normally cools it.

Rubber is, however, anomalous, since it warms when it is stretched, as can easily be tested by suddenly stretching an elastic band placed in contact with the lips.

§ 126. Entropy

If a body *absorbs* a quantity of heat Q at a temperature T it is said to have *increased* its entropy by an amount $\frac{Q}{T}$.

This statement is essentially our definition of the term entropy, since we are usually only concerned with changes of entropy.

In the Carnot cycle which has been discussed, a quantity of heat Q_I was absorbed at a temperature of T_1, so that the entropy increased by an amount $\frac{Q_1}{T_1}$. On the other isothermal, a quantity Q_2 of heat was rejected at the temperature T_2, so that the entropy here increased by an amount $\left(-\frac{Q_2}{T_2}\right)$. The total increase in entropy in the Carnot cycle was, therefore,

$$\left(\frac{Q_1}{T_1}\right) - \left(\frac{Q_2}{T_2}\right),$$

which must be zero, since

$$\text{efficiency} = \frac{Q_1 - Q_2}{Q_1} = \frac{T_1 - T_2}{T_1},$$

and

$$\frac{Q_2}{Q_1} = \frac{T_2}{T_1}.$$

This theorem of the simple Carnot cycle can now be extended to cover *any* reversible cycle by considering it as being built up of an infinite number of separate Carnot cycles of infinitesimally small size. For each of these separate Carnot cycles the change in entropy is zero, so that for the whole of the reversible cycle the total change in entropy will also be zero (Fig. 163).
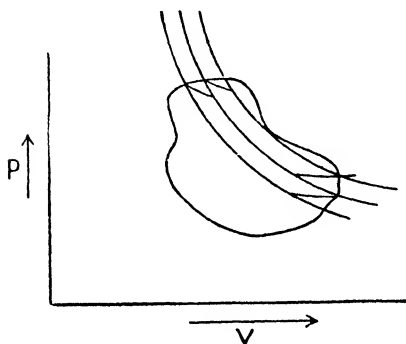
Fig. 163

Any practical heat engine will take in some of its heat at a temperature lower than that at which it is produced in the furnace, i.e. at a temperature less than T_1. In consequence the entropy will increase by more than $\frac{Q_1}{T_1}$, so that, over the whole cycle, the entropy will register an increase. This is an expression of the practical statement that in all changes which are not perfectly reversible the total entropy of the system tends to increase.

§ 127. Thermoelectric Currents

Thermoelectricity is the most usual example of the application of this fact that the change of entropy of a perfect reversible system is zero.

Consider a closed electrical circuit of two wires A and B, as in Fig. 164. The wire junctions are maintained respectively at the temperatures T and $T - dT$.

There appear to be two entirely distinct thermoelectric effects occurring in such circuits. The first, called the Peltier effect, is such that a quantity of heat Pq is absorbed at the hotter junction when a quantity of electricity q passes round

the circuit. At the cooler junction P, the Peltier co-efficient, will be somewhat different and can be represented as

$$P - \left(\frac{dP}{dT}\right)\delta T.$$

At this junction, however, since the current is now proceeding in the reverse, or negative, direction, heat will be emitted and not absorbed.

In addition to this Peltier effect, there appears also to exist another reversible effect called the Thomson effect, which we may repre-sent quantita-tively by σ_a and σ_b for the two wires A and B respectively. The nature of this re-

Wire A

T   T−dT

Wire B
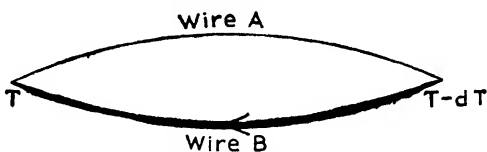
FIG. 164

versible effect is such that if a quantity of electricity q is taken from the hot junction (T) to the cold one (T − δT) along the wire A, a quantity of heat $q\sigma_a \delta T$ will be transferred from the hot to the cold junction. In a similar way, if the electricity is taken along the wire B *from the cold to the hot junction* the quantity of heat transferred from *the hot to the cold junction would be* −$\sigma_b qdT$.

Combining the two effects we see that the heat which is absorbed and which is not re-emitted is

$$Pq - \left(P - \frac{dP}{dT}\delta T\right)q + \sigma_a q\delta T - \sigma_b q\delta T.$$

This heat must correspond to the electrical work done, and so

$$\delta Eq = Pq - \left(P - \frac{dP}{dT}\delta T\right)q + \sigma_a q\delta T - \sigma_b q\delta T,$$

or

$$\delta E = \frac{dP}{dT}\delta T + (\sigma_a - \sigma_b)\,\delta T.$$

If, in this equation, δE is to be measured in volts the quantities P and σ should be expressed in joules (10^7 ergs) per coulomb.

All the heat interchanges which occur as the electricity, q, is taken round the above cycle are reversible, if we ignore the

K

normal heating C^2R of the current. It does seem theoretically justifiable to ignore this because the reversible effects we have just been considering depend only on q, and do not depend on the speed with which we take q round the circuit. Thus, if we allow the electricity q to move infinitely slowly the current will be zero, so that the total non-reversible heating effect due to the C^2R term will also be zero. (If C tends to zero, and t to infinity, Ct may be finite although C^2t tends to zero.)

It is therefore legitimate to regard the above cycle as thermodynamically reversible so that it is possible to apply our equation that

$$\text{efficiency} = \frac{Q_1 - Q_2}{Q_1} = \frac{T_1 - T_2}{T_1},$$

either in this form or, as is here more convenient, as the deduction from it, that the total entropy change $\frac{dQ}{T}$ in a reversible cycle is zero.

At the junction T, the entropy increase is $\frac{Pq}{T}$.

Therefore at the junction $T - dT$, the increase of entropy will be

$$- \left\{ \frac{Pq}{T} - \frac{d}{dT}\left(\frac{Pq}{T}\right)\delta T \right\}.$$

Along the wire A the increase of entropy will be

$$\frac{\sigma_a\, q\delta T}{T},$$

and along the wire B it will be

$$\frac{-\sigma_b\, q\delta T}{T}.$$

The total increase in entropy is zero, and so

$$\frac{Pq}{T} - \left\{ \frac{Pq}{T} - \frac{d}{dT}\left(\frac{Pq}{T}\right)\delta T \right\} + \left(\sigma_a - \sigma_b \right)\frac{q\delta T}{T} = 0.$$

$$\therefore \sigma_a - \sigma_b = -T\frac{d}{dT}\left(\frac{P}{T}\right) = -T\left\{ \frac{1}{T}\frac{dP}{dT} - \frac{P}{T^2} \right\}.$$

$$\therefore \sigma_a - \sigma_b = -\left(\frac{dP}{dT} - \frac{P}{T} \right).$$

Combining this result with that previously found, viz.,

$$\delta E = \frac{dP}{dT}\delta T + (\sigma_a - \sigma_b)\delta T,$$

we see that $\dfrac{\delta E}{\delta T} = \dfrac{P}{T}$,

or that $P = T\dfrac{dE}{dT}.$

The really difficult part to remember in this theory is the method of expressing the entropy at $(T - \delta T)$ as

$$- \left\{ \frac{Pq}{T} - \frac{d}{dT}\left(\frac{Pq}{T}\right)\delta T \right\},$$

so the student is warned to be on his guard against forgetting this point.

§ 128. Experimental Determination of the Thermoelectric Coefficients

It is easy to obtain in the laboratory a graph connecting the thermoelectric E.M.F. and the temperature for any pair of materials. Since this graph is usually a parabola it is convenient to obtain a straight line by plotting $\dfrac{dE}{dT}$ against T.

Occasionally this graph is obtained by direct experiment by heating the two baths in which the thermoelectric junctions are immersed, in such a way that the difference in temperature between them is accurately constant and is small. This gives directly the mean value of $\dfrac{dE}{dT}$ over the small range of temperature considered.

The area LMPQ in Fig. 165 clearly represents the total E.M.F. observed experimentally between the two temperatures T_1 and T_2 because it is

$$\int \frac{dE}{dT}\delta T.$$

The area of the rectangle OPQS represents the total Peltier effect of the temperature T_1 if the graph has been drawn with its origin corresponding to zero temperature on the gas or absolute scale. This follows because we have shown that
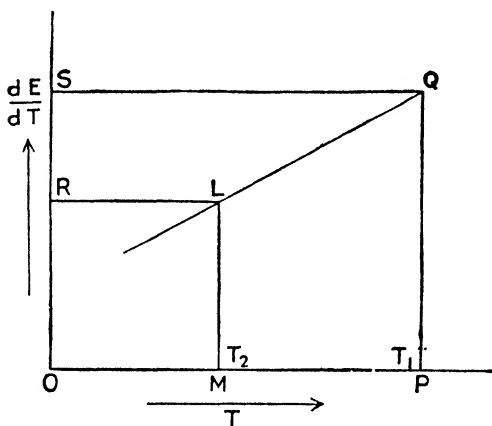
$$P = T \frac{dE}{dT}.$$

FIG. 165

Similarly, the Peltier effect at the temperature T_2 is represented by the area OMLR.

We can also show that the area RLQS represents the Thomson effect for unit quantity. This area is essentially

$$\int T \times \left(\text{small change in } \frac{dE}{dT}\right).$$

But

$$\frac{dE}{dT} = \frac{P}{T},$$

and so area RLQS $= \int T \times \left(\text{small change in } \frac{P}{T}\right).$

Now we have already proved that

$$(\sigma_a - \sigma_b) = -T\frac{d}{dT}\left(\frac{P}{T}\right).$$

$\therefore$ area RLQS $= \int(\sigma_a - \sigma_b)dT,$

which is the total Thomson effect for unit quantity of electricity between these two temperature limits.

CHAPTER XII

LIQUEFACTION AND LOW TEMPERATURES

§ 129. The Cascade Process.   § 130. The Claude Method.   § 131. The Joule-Kelvin Effect.   § 132. The Numerical Value of the Joule-Kelvin Effect.   § 133. The Linde Process.   § 134. The Liquefaction of Hydrogen and Helium.   § 135. The Thermal Insulation of Liquefied Gases. § 136. The Measurement of Low Temperatures.   § 137. Liquid and Solid Helium.   § 138. Super-conductivity.   § 139. Cooling by Adiabatic Demagnetization.

§ 129. The Cascade Process

It is well known that there are three processes by which gases may be liquefied.   In the Pictet or cascade process, the low temperature is reached by successive stages; methyl chloride is first liquefied by compression at ordinary temperatures, and the liquid is then made to evaporate rapidly by reducing the pressure above it with powerful pumps.   At the temperatures that can be reached in this way ethylene can be liquefied by compression.   This, in turn, is evaporated rapidly, producing a temperature at which oxygen can be liquefied by pressure.

The limit with this process is reached with nitrogen, because no substance can be found which remains liquid or has an appreciable vapour pressure at temperatures as low as the critical point of hydrogen.

Students should refer to the ordinary text-books for diagrams and more complete details of this process.

§ 130. The Claude Method

In the Claude method for liquefaction, compressed gas is allowed to expand, doing external work in addition to the

internal work done against its internal molecular attractions. A familiar example of this method of cooling occurs when a 'sparklet soda-water' bulb is pierced; the compressed gas expands into the syphon and the bulb gets so cold that hoar frost frequently forms.

A diagrammatic form of the apparatus used in the Claude method is shown in Fig. 166. When the gas has entered the
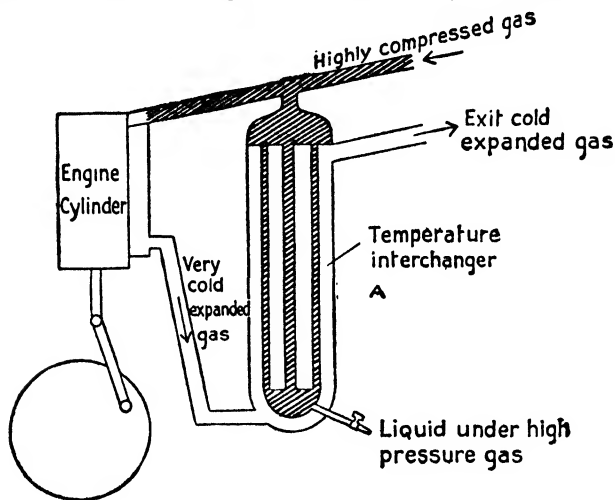
FIG. 166

cylinder and the valve has closed, it expands adiabatically, doing work on the piston which may be used to help to drive the compressor.

If we assume that the expansion is adiabatic,

$$p_1 v_1{}^\gamma = p_2 v_2{}^\gamma,$$

and, since $\qquad p_1 v_1 = RT_1,$ and $p_2 v_2 = RT_2,$

it follows that $\qquad \left(\dfrac{p_1}{p_2}\right)^{\frac{\gamma-1}{\gamma}} = \dfrac{T_1}{T_2},$

so that an expansion from 100 to 10 atmospheres should change the temperature from 300° A. to 155·6° A. On account of the

frictional heat produced between the piston and the cylinder, and the fact that the expansion cannot be completely adiabatic, the actual temperature produced in any single expansion is not nearly so low as the theoretical.

The cooled gas, after its exit from the cylinder, passes through a heat interchanger and cools the incoming gas. Finally, when the temperature is low enough, the high pressure gas in the condenser A (Fig. 166) will liquefy and can be drawn off at the bottom.

One great difficulty of the Claude process is in finding a suitable cylinder lubricant which is not solid at the low temperatures prevailing. Once the gas has begun to liquefy in the cylinder the liquid itself acts as a lubricant.

Although this Claude apparatus has been constructed to operate successfully, most liquefaction plants employ the Linde process, which completely obviates the lubrication problem and the necessity of the complicated expansion cylinder and piston, at the expense of only a slight loss of efficiency—as will be shown later.

§ 131. The Joule-Kelvin Effect

The basis of the normal commercial—or Linde—process for the liquefaction of gases was the discovery by Joule and by Joule-Kelvin of the cooling that may occur when a compressed gas is allowed to expand through a fine jet without performing external work. The effect may now be examined in detail.

If the high-pressure gas in the cylinder A of Fig. 167
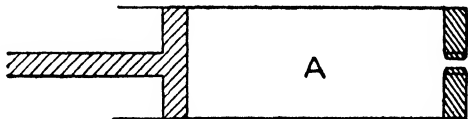
FIG. 167

is allowed to expand adiabatically into the atmosphere through the small hole, without any movement of the piston, it will be doing external work and will cool. Since every cubic centimetre of the gas has expanded by the same fraction, the fall of temperature will be spread equally over the whole

volume of the gas, and will not, for example, be concentrated near the hole.

Similarly, if the hole were closed and the piston moved forward work would be done on the gas and it would heat up uniformly throughout its whole volume.

If the two operations were performed simultaneously in such a way that the pressure of the gas in the cylinder remained unchanged, there would be no change in temperature of the gas in the container. In the process the gas emerging from the hole has compressed the atmosphere and has, therefore, increased the total heat energy of the atmosphere by an amount which is equivalent to the work done by the piston.

Now consider the Joule-Kelvin experiment, in which the gas
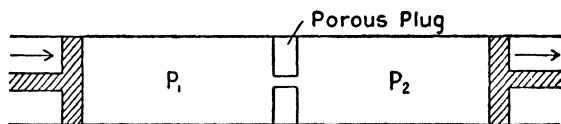
FIG. 168

was made to pass through a porous plug from a cylinder and piston on the left to another on the right (Fig. 168). The holes in the porous plug were sufficiently fine and numerous to permit a convenient difference of pressure between the two cylinders with reasonable speeds of gas flow; and the speeds of the two pistons were adjusted so that the pump on the right moved so much faster than that on the left that it exactly compensated for the change in volume of the gas as its pressure changed from p_1 to p_2.

We have seen above that the temperature of the gas in the first cylinder must remain constant if the pressure p_1 is constant.

The work done by the gas entering the second cylinder is clearly p_1v_1, while the gas in this cylinder does work p_2v_2 on the second piston. The temperature of the gas in this second cylinder will, therefore, fall by an amount corresponding to

$$(p_2v_2 - p_1v_1),$$

which will, of course, be zero, if Boyle's law is obeyed accurately.*

Ordinary gases obey Van der Waal's equation approximately, and so have attractive forces between molecules equal to $\frac{a}{v^2}$.

In the above expansion, work must have been done *by the gas* against these attractive forces equal to

$$\int_{v_1}^{v_2} \frac{a}{v^2} dv = \left(-\frac{a}{v_2} + \frac{a}{v_1} \right).$$

The temperature of the gas should therefore fall by an amount corresponding to the total *net* work done:

$$(p_2 v_2 - p_1 v_1) + \left(-\frac{a}{v_2} + \frac{a}{v_1} \right).$$

This fall in temperature of the gas in the second cylinder is, as we have shown, made up of two parts; the one is due to external work done as a result of a departure of the gas from Boyle's law, while the other is due to the internal molecular attractions. When a gas expands and has v_2 larger than v_1, work must always be done *against* these molecular attractions so that the gas must cool. On the other hand, the departure from Boyle's law may lead to either an increase or a decrease

---

* This same fact that a perfect gas does not cool when it is allowed to expand into another enclosure without driving a piston back and thus doing external work, may be shown if we consider the dynamical mechanism by which a gas contained in a cylinder and piston does cool when the piston is moved back. Thus, since the coefficient of restitution for molecules hitting anything is unity, we have:

$$v_1 - v_2 = -(u_1 - u_2).$$

So,
$$v_1 = -u_1 + (u_2 + v_2).$$

The molecule therefore rebounds with a smaller speed from a surface which is moving away from it; it has therefore a smaller kinetic energy, and so is at a lower temperature.

In Fig. 168, however, a molecule which hits the solid wall near the aperture will be returned with an undiminished average energy and unchanged temperature. Similarly any molecule which is moving through the aperture will (possibly after hitting other gas atoms) eventually hit the distant wall. Then because this wall is stationary it will be returned with an unchanged average velocity and temperature

The gas therefore cools if it hits a moving piston and does not cool if it passes through an aperture on to a stationary wall.

*K

in temperature on free expansion. This is shown by the experimental graphs of pv against pressure, indicating that at low temperature pv first diminishes and then increases with increasing pressure. Thus gas at low temperatures and pressures has p_2v_2 bigger than p_1v_1, and will cool, because of the departure from Boyle's law. Conversely, if the temperatures or pressures are high, p_1v_1 will be bigger than p_2v_2, and the gas will heat up because of the departure from Boyle's law.

The total resultant of the internal work and of the external work done in virtue of this departure from Boyle's law, will depend essentially on the relative magnitudes of the two effects. However, it appears in general, that at high temperatures and pressures p_1v_1 is so much bigger than p_2v_2 that the resultant effect is an increase in temperature of the expanded gas, whereas at low temperatures the reverse occurs.

§ 132. The Numerical Value of the Joule-Kelvin Effect

Numerical values for the cooling effect (in ergs) of a gas may be obtained if we assume that Van der Waal's equation is obeyed exactly. Then

$$\left(p + \frac{a}{v^2}\right)\left(v - b\right) = \mathrm{RT},$$

$$pv + \frac{a}{v} - pb - \frac{ab}{v^2} = \mathrm{RT},$$

$$\left(pv - \frac{a}{v}\right) = \mathrm{RT} + \frac{ab}{v^2} + pb - \frac{2a}{v},$$

so that, if we neglect the term $\frac{ab}{v^2}$ because it is of the second order of small quantities, the total work done in the Joule-Kelvin effect will be:

$$\left(p_2v_2 - \frac{a}{v_2}\right) - \left(p_1v_1 - \frac{a}{v_1}\right) = \left(p_2 - p_1\right)b - 2a\left(\frac{1}{v_1} - \frac{1}{v_2}\right)$$

$$= \left(p_1 - p_2\right)\left(\frac{2a}{\mathrm{RT}} - b\right),$$

if we substitute the approximate value $\dfrac{p}{RT}$ for $\dfrac{1}{v}$ in the terms

$\dfrac{a}{v_1}$ and $\dfrac{a}{v_2}$.

If the gas is expanding so that p_1 is greater than p_2 the gas will cool if $\left(\dfrac{2a}{RT}\right)$ is greater than b.

The inversion temperature, at which the cooling effect is zero, and above which it is negative, is given by the relation

$$\frac{2a}{RT} - b = 0.$$

It is interesting to compare the magnitude of this Joule-Kelvin effect with the external work done when a perfect gas expands; this external work done may be written as

$$\int p\,dv = RT\int \frac{dv}{v} = RT\,\log_e \frac{v_2}{v_1} = RT\,\log_e \frac{p_1}{p_2}.$$

Notice that the work done will always be expressed in ergs if the pressure is in dynes per sq. cm., the volume in cc., and if the values of R, a, and b are calculated using these units of pressure and volume.

TABLE X. DATA FOR CALCULATION OF COOLING EFFECT
WITH OXYGEN

For 1 gm. of oxygen, if the pressures are in dynes per sq. cm. and the volume is in cc.	R $=$ $2 \cdot 59 \times 10^6$ ergs per A.° a $=$ $1 \cdot 36 \times 10^9$ dynes cm.^{4} b $=$ $\cdot 995$ cc.
1 atmosphere pressure	$= 1 \cdot 013 \times 10^6$ dynes per sq. cm.
	$C_v = 0 \cdot 173$ cals. per gm. per A.°
	$C_p = 0 \cdot 242$ cals. per gm. per A.°
External work done by perfect gas	$= RT\,\log_e \left(\dfrac{p_1}{p_2}\right)$ ergs
Work done in Joule-Kelvin effect	$= \left(\dfrac{2a}{RT} - b\right)\left(p_1 - p_2\right)$

From the data in Table X we find that:

An expansion from 2 to 1 atmospheres pressure at 273° A.

The external work done by a perfect gas $= 4\cdot91 \times 10^8$ ergs.

The internal 'Joule-Kelvin' work done $= 2\cdot91 \times 10^6$ ergs.

An expansion from 2 to 1 atmospheres pressure at 173° A.

The external work done by a perfect gas $= 3\cdot1 \quad \times 10^8$ ergs.

The internal 'Joule-Kelvin' work done $= 5\cdot11 \times 10^6$ ergs.

An expansion from 200 to 20 atmospheres pressure at 273° A.

The external work done by a perfect gas $= 1\cdot63 \times 10^9$ ergs.

The internal 'Joule-Kelvin' work done $= 5\cdot24 \times 10^8$ ergs.

An expansion from 200 to 20 atmospheres pressure at 173° A.

The external work done by a perfect gas $= 1\cdot03 \times 10^9$ ergs.

The internal 'Joule-Kelvin' work done $= \quad \cdot92 \times 10^9$ ergs.

These figures show that the Joule-Kelvin effect becomes of greater importance when the molecules become closer together either through the application of great pressure or of low temperature, and that, under the conditions prevailing in commercial liquid air machines, the cooling by the Joule-Kelvin effect is of much the same magnitude as that produced by the external work done during expansion by a perfect gas.

§ 133. The Linde Process

Air is liquefied in the Linde regenerative process by expanding it through a valve from about 200 to 20 atmospheres pressure, so that as a consequence of the Joule-Kelvin effect it is cooled by approximately 40° C. This cooled expanded gas is used to cool the incident compressed gas, so that the gas arriving at the compression valve gets progressively cooler until it ultimately liquefies.

The gas entering the machine must be purified very carefully from water vapour and from carbon dioxide, as these gases would solidify and clog the valve. The compression is frequently carried out in two or more stages, the low pressure pumps compressing to about 20 atmospheres only; this is

one of the reasons why the gas is not generally expanded to a lower pressure than 20 atmospheres. After each compression, the gas is water-cooled to normal room temperatures before being cooled by the expanded gas. The heat interchanger, where the cooled expanded air cools the incident compressed air, generally consists of two concentric tubes several hundreds of yards long, which are coiled spirally to economize space. They must be well lagged.

Although the Joule-Kelvin cooling between these pressure ranges is only approximately one-half of that occurring in a Claude expansion, the latter process is very rarely employed. This is due partly to the difficulty of lubrication of the Claude piston and partly to the fact that the temperature of the piston and cylinder must tend to change rapidly from about $-30°$ C. to $-182°$ C., whereas in the Linde process there is a continuous gas flow in which the metal parts of the apparatus are at stationary temperatures. This last point is one of the utmost practical importance, and the student is advised to obtain a clear-cut picture of the process. First comes the initial pumping system to deliver gas at about 20 atmospheres pressure to the liquefier proper. · Next there is the high-pressure compressor and then the long high-pressure pipe leading to the valve. The first part of this pipe is at the relatively high temperature ($\backsim$ 250° C.) produced by the compression; further along the pipe the temperature has dropped to about 15° C. because of the water-cooling. The high-pressure pipe then enters the middle of its surrounding low-pressure one. Some thirty feet further on the temperature of the pipe will be perhaps ·0° C., after another thirty feet $-15°$ C., and so on. Once liquefaction has begun the temperature of the pipes will remain quite constant, one end being at $-180°$ C. and the other at $+250°$ C.

§ 134. The Liquefaction of Hydrogen and Helium

The normal room temperatures are materially higher than the inversion temperatures of hydrogen and helium, and these gases must be cooled to about 190° A. and 25° A. respectively

before any appreciable cooling occurs with the Joule-Kelvin effect.

To liquefy hydrogen it is therefore necessary to build first a liquid air plant, to erect a hydrogen compressor, a water-cooler, and a liquid air cooler, and then to carry out the normal Linde process with this 'liquid air cooled' hydrogen.

Large-scale liquid helium machines are even more complicated, because it is necessary here to cool the gas first with liquid air and then with *liquid* hydrogen which is boiling under reduced pressure.

§ 135. The Thermal Insulation of Liquefied Gases

Liquid air is generally kept in small quantities in silvered, double-walled Dewar flasks with an evacuated annular space. Such containers must never be corked, but should instead be closed with a loose pad of cotton-wool, since the increasing pressure of the evaporating gas will inevitably burst the strongest closed container.

The heat received per second—and therefore the mass of gas evaporating—is proportional to the surface area of the inner cold container. It is therefore most economical to use a spherical shape, for which the ratio of volume contained to surface area is a maximum. Similarly, the *percentage* loss by evaporation will vary inversely as the linear dimension of the container, because the mass evaporating is proportional to $4\pi r^2$, while the quantity present is proportional to $\frac{4}{3}\pi r^3$.

It is interesting to note that, in Germany, railway or motor containers of several tons capacity are made to transport liquid oxygen for hundreds of miles from a few very large producing centres to a much larger number of sub-stations, where the liquid is evaporated to fill the oxygen cylinders normally used in commerce. The advantage of this method is that it is much more economical to produce oxygen by the fractional distillation of liquid air in works of large than of small capacity; but this was not previously done because of the great expense of transporting heavy cylinders capable of holding *gas* at 120 atmospheres.

In several parts of the world natural gas escaping from oil-wells, which was formerly run to waste, is now liquefied and carried hundreds of miles in ships, as a liquid, before it is finally evaporated directly into the main gas works of large towns. The ships are most conveniently driven by gas-engines using the natural evaporation of the liquefied gas.

The thermal insulation of liquid hydrogen is a much more difficult problem than that of liquid air, since its temperature and its latent heat are so much lower. The Dewar flask containing the liquid hydrogen is generally kept in another flask filled with liquid air. We know that the major heat losses from Dewar flasks are due to radiation, which obeys Stefan's fourth power law approximately. Thus, since the temperatures of liquid air and of liquid hydrogen are respectively 75° A. and 20° A., the heat lost from hydrogen without a liquid air container is $\propto (273)^4 - (20)^4 \propto 5 \cdot 5 \times 10^9$, and the heat lost from hydrogen with a liquid air container $\propto (75)^4 - (20)^4 \propto 3 \cdot 1 \times 10^7$. Thus the radiation losses from hydrogen are reduced by the liquid air bath to one-hundredth of their previous value.

Heavy Copper Connecting Lug

Liquid H_2 Liquid Air

Dewar Flasks
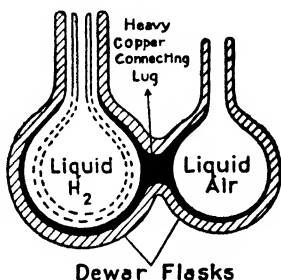
Fig. 169

Kapitza avoided the necessity for using large baths of liquid air by the method shown in Fig. 169. This method is particularly efficient because, at the temperature of liquid air, the thermal conductivity of copper is very great.

Liquid helium is generally kept in a triple Dewar flask, surrounded respectively by liquid hydrogen and liquid air. In order to see the liquid helium it is necessary that none of the vacuum vessels be silvered. The final vessel containing liquid air is particularly convenient, because otherwise solidified air will be deposited on the outside of the hydrogen container and will make its surface cloudy, so that the helium cannot be seen.

§ 136. The Measurement of Low Temperatures

The ultimate ideal is to measure all temperatures on the thermodynamical, work, or absolute scale of temperature. This ideal scale is very nearly the same as the practical gas scale, and it is possible to correct practical gas scale temperatures to absolute values

The gas scale is, therefore, used as the practical standard, and certain basic fixed points have been determined on the scale down to the boiling-point of liquid oxygen at ordinary pressures ($-182 \cdot 97°$ C.).

For lower temperatures, first hydrogen and then helium thermometers may be used. At temperatures below the liquefying point of helium it is possible to use the helium thermometer, providing the pressure of the gas in it is kept so low that it is, throughout the range concerned, below its vapour pressure.

Temperatures from $1 \cdot 2°$ A. $\rightarrow 0 \cdot 7°$ A. cannot be determined by gas thermometers, but are guessed at by extrapolating the vapour-pressure curve or the resistance-temperature relation. Aid in the above extrapolation is to be had by an analogy with the vapour-pressure curve of hydrogen, using the law of corresponding states, while there is now a reasonably accurate theory of the variation of normal electrical resistance with temperature near the absolute zero.

Temperatures below $0 \cdot 7°$ A. are measured by employing Curie's law, according to which the magnetic susceptibility of a paramagnetic substance is inversely proportional to its absolute temperature. The sensitivity of the method therefore increases as the temperature decreases. As an example, we may notice that at $1°$ A. the molecular susceptibility of iron ammonium alum is $\backsim 4$, which may be compared with the value ($\backsim 1,000$) obtained with iron at ordinary temperatures. This method of measuring temperatures by investigating magnetic susceptibilities has been used for temperatures as low as $0 \cdot 015°$ A. where the molecular susceptibility is of the order of 240.

Once an adequate series of fixed points has been determined with a gas thermometer it is possible to measure temperatures much more conveniently by means of resistance thermometers. In the temperature range —190° C. to 0° C., it is usual to employ a standard platinum resistance thermometer which is calibrated by reference to the internationally agreed fixed points.

Below a temperature of —190° C. or 83° A. platinum becomes an unsuitable material for a resistance thermometer, as its resistance then varies very slowly indeed with temperature.

Lead seems particularly convenient from about 80° A. to 20° A., while, for temperatures below that, manganin or unannealed phosphor-bronze is employed. At the temperature of liquid helium this latter material—non-annealed phosphor-bronze—shows a change of resistance with temperature which is fifty-five times as great as that shown by constantan.

§ 137. Liquid and Solid Helium

Liquid helium, obtained in the way already indicated, is a colourless, mobile liquid. Its density—which is approximately one-eighth of that of water—varies markedly with temperature and pressure, and it shows a definite temperature of maximum density, which varies markedly with pressure.

Table XI shows the general manner in which the vapour pressure of helium varies with the temperature. Graphs derived from these figures may be employed as a very convenient and accurate practical means of estimating the corresponding temperatures.

TABLE XI

Temperature A.°	4·22°	2·3°	1·54°	1·14°	0·960°	0·844°
Vapour Pressure in cm. Hg	76·0	5·00	0·500	0·05	0·0100	0·0025

If liquid helium is made to boil under reduced pressure by evacuating the space above it with powerful pumps, the latent heat of vaporization will be absorbed from the liquid
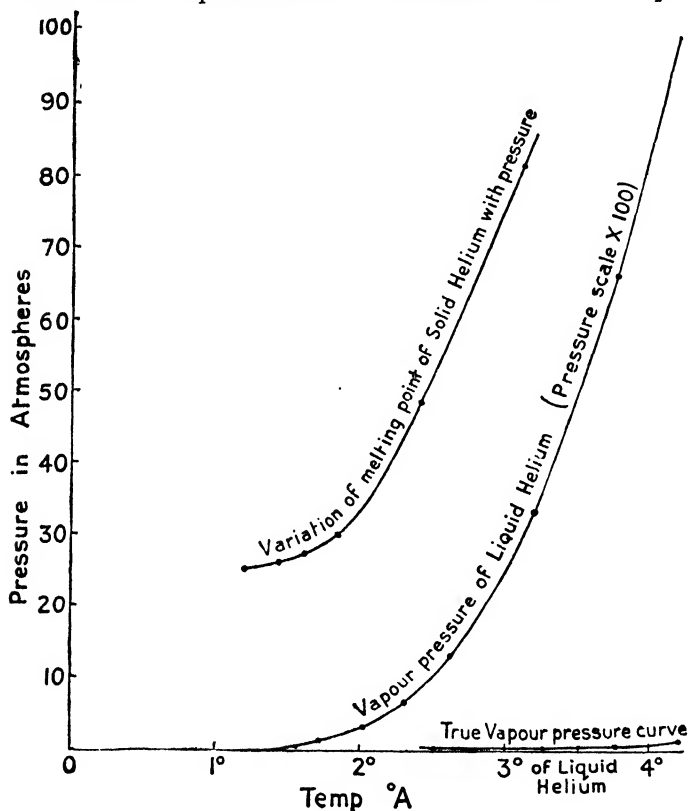
FIG. 170

itself and the temperature will fall until the vapour pressure is equal to the best vacuum that the pump can maintain. In this way Keesom reached an estimated temperature of 0·71° A. This temperature must be very near the limit which can be reached by such means, as can be seen by

inspecting the above vapour pressure table. In 1933, however, a new method of cooling was developed by which temperatures as low as 0·005° A. have been reached. This method will be discussed later.

Kamerlingh Onnes succeeded in reaching temperatures of 0·7–0·8° A., but never obtained any trace of solidification. Keesom argued that, since the melting-point of solid hydrogen was greatly increased by the application of pressure, it might be possible that well-cooled liquid helium might also solidify if the external pressure was great enough. On 21st June 1926 he therefore cooled liquid helium to 1·10° A. by the usual method, and then applied a pressure of 200 atmospheres. Although no visible difference was produced in the helium, yet he showed that the helium had really solidified by trying to move with a magnet a small piece of iron in the liquid. The iron fragment never appeared completely embedded in the liquid because the light used to illuminate it always kept the liquid in the immediate vicinity of the iron rather warmer than the surrounding medium.

Keesom also determined the melting-point of solid helium at different pressures, and showed that the graph curves at the lowest temperatures so as to become more and more nearly parallel to the temperature axis.

Fig. 170 shows no tendency for this curve to intersect the vapour-pressure curve at the triple point, as is usual with all other substances. Thus solid and gas cannot exist together, and the solid always melts before it vaporizes at all temperatures.

§ 138. Superconductivity

The variation with temperature of the electrical resistance of metals was studied by Dewar and Fleming in the temperature range of liquid air and liquid hydrogen as early as 1893. Their results suggested that the electrical resistance of pure metals would be zero at the absolute zero of temperature.

In addition, almost any theories of the electrical resistance

of metals lead to the same conclusion that the resistance should decrease with a decrease in temperature and should tend to zero as the temperature tends to the absolute zero.

Both the theories and the experimental results agreed, therefore, in expecting that the resistance/temperature relation would curve *smoothly* to a zero value of resistance, at the absolute zero of temperature. In 1911, however, Kamerlingh Onnes discovered that the electrical resistance of mercury fell sharply at 4·2° to zero. Since there is no measurable resistance, an electric current started in a closed ring of the substance will continue undiminished in intensity until the metal is heated again above this temperature. This heating must be performed by some external source and cannot be produced by the current itself, since the resistance is zero.

Other metals have since been shown to pass into this superconducting state, as it is called; the transition temperatures are shown in Table XII.

TABLE XII. TEMPERATURES AT WHICH METALS BECOME
SUPERCONDUCTIVE

Aluminium 1·14° A.	Titanium 1·75° A.	Vanadium 4·3° A.	Zinc 0·79° A.	Gallium 1·05° A.
Niobium 9·2° A.	Cadmium 0·6° A.	Indium 3·37° A.	Tin 3·7° A.	Tantalum 4·4° A.
Mercury 4·22° A.	Thallium 2·37° A.		Lead 7·2° A.	Thorium 1·5° A.

Some alloys and chemical compounds also become superconducting, even though, in certain examples, their constituents are not known to be superconductors.

The experimental investigation of superconductivity is generally made by taking a closed ring of the metal immersed in liquid helium and placing it between the poles of a very powerful electro-magnet. The pressure on the helium is then reduced, it boils, and the temperature falls. At frequent

intervals the field of the electro-magnet is removed so that an E.M.F. equal to $\dfrac{dN}{dt}$ is induced in the ring. If the metal is superconducting, a large current will flow which will produce a magnetic field able to be detected by a simple compass.

On a first consideration one usually imagines that the induced E.M.F. must produce an infinite current in a wire of zero resistance. This cannot, however, occur, because the total change in the lines of induction threading the ring is the change in the number produced by the electro-magnet plus the change in the number produced by the ring itself. The current in the ring, therefore, only rises to such a value that the total magnetic flux through it is the same as it was before the field of the electro-magnet was reduced.

One complication of this description is that the application of a magnetic field causes superconductivity to appear in a metal at a lower temperature than normally. Thus, with mercury, the superconducting transition temperature is 2·2° A. in a field of 300 gauss, and 3·7° A. in a field of 100 gauss. If, therefore, we were to cool a mercury ring in a field of 300 gauss to 3·7° A. and then suddenly remove the external field, the current in the ring would rise to a value corresponding to 100 gauss only.

§ 139. Cooling by Adiabatic Demagnetization

We have already seen that Kamerlingh Onnes reached his lowest temperature of 0·7–0·8° A. by evaporating liquid helium with exceedingly powerful pumps. This temperature seems very near to the lower limit which could be obtained by such methods, because the vapour pressure of liquid helium is then decreasing very rapidly indeed with decrease in temperature, and is only $2·5 \times 10^{-3}$ cm. Hg at 0·84° A.

Debye and also Giauque suggested in 1926 that it might be possible to use a new method by which cooling is effected by 'adiabatic demagnetization.' This method has since been

tested, and has been found to be so efficient that temperatures as low as 0·005° A. have now been reached.

The basic idea can be described roughly in simple terms by pointing out that the 'molecular magnets' of a paramagnetic body tend to form closed rings of the type shown in Fig. 171 (a). If the substance is now placed in a very powerful magnetic field the 'molecular magnets' will all arrange themselves along the lines of force as shown in Fig. 171 (b). When
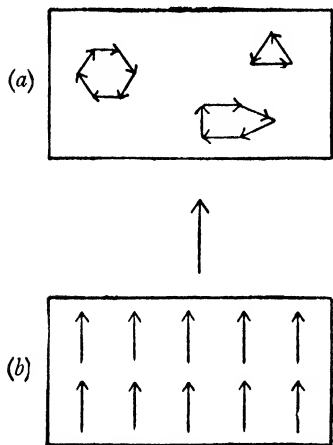
FIG. 171

the system is in this latter state the temperature is reduced to the lowest convenient value (∽ 1° A.) with evaporating liquid helium. All trace of residual gas is then removed from the space round the magnetic specimen, so that there is nothing present (except very fine supporting wires) to conduct heat.

The magnetic field is then almost completely removed and the 'molecular magnets' immediately swing round into their closed chains under their mutual magnetic actions. Since the magnetic system is thermally insulated and the demagnetization is adiabatic, the energy needed for this can only come from the kinetic energy of the molecules, so that the temperature must fall.*

* This cooling which can be produced by adiabatic demagnetization can also be readily explained in a general way by a consideration involving entropy. It is well known that entropy is a measure of disorder; and that the thermal change of entropy of a system $\frac{dQ}{T}$ is a measure of the change of disorder in respect of velocity. Now if a magnetic field is applied to a paramagnetic substance, the elementary magnets become completely ordered in respect of direction, so that the thermal order must decrease, which means that the temperature must rise.

This method is much more efficient than would otherwise be expected because, at these very low temperatures, specific heats have values which are very much smaller than their normal values at room temperatures. Thus the specific heat of tin at 273° A. is 0·11 calories per gm. per degree Centigrade, while at 3·8° A. it is only 0·000047. Moreover, at these low temperatures the magnetic susceptibility of the salts employed (cerium fluoride, iron ammonium alum, etc.) is very great.

The measurement of these low temperatures is usually effected by determining experimentally the mechanical force of attraction between the salt employed and the residual magnetic field left when the current in the electro-magnet is switched off. From this mechanical force of attraction the susceptibility is calculated, and from the susceptibility, in turn, the temperature is determined with the help of Curie's theoretical relation, which appears certain to be at least approximately correct.

Correspondingly, if the substance is now cooled thermally again by the evaporation of liquid helium, is insulated thermally by removing the surrounding gas, and is then demagnetized, the thermal order must increase—which means that the temperature must fall.

There is no really new idea involved in this cooling by adiabatic demagnetization. When iron is taken round a hysteresis cycle the work done is $=\int H dI$ per unit volume. Thus the iron heats when it is magnetized and, to a less extent, cools when demagnetized. The difference between the ordinary hysteresis experiments and those at low temperatures is one of degree and not of kind. The field is large, $\sim$ 30,000 gauss, the susceptibility high, and the specific heat low at these low temperatures.

INDEX

INDEX